PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68 — A2

(11) International Publication Number: **WO 00/12760**

(43) International Publication Date: 9 March 2000 (09.03.00)

(21) International Application Number: PCT/US99/19768

(22) International Filing Date: 27 August 1999 (27.08.99)

(30) Priority Data:

09/141,825	28 August 1998 (28.08.98)	US
09/172,711	13 October 1998 (13.10.98)	US
09/172,108	13 October 1998 (13.10.98)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	09/141,825 (CIP)
Filed on	28 August 1998 (28.08.98)
US	09/172,711 (CIP)
Filed on	13 October 1998 (13.10.98)
US	09/172,108 (CIP)
Filed on	13 October 1998 (13.10.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CUNNINGHAM, Mary, Jane [US/US]; 1204 Manet Drive, Sunnyvale, CA 94087 (US). ZWEIGER, Gary, B. [US/US]; 2700 Del Medio Court, #124, Mountain View, CA 95023 (US). PANZER, Scott, R. [US/US]; 965 East El Camino, #621, Sunnyvale, CA 94087 (US). SEILHAMER, Jeffrey, J. [US/US]; 12555 La Cresta, Los Altos, CA 94022 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: TOXICOLOGICAL RESPONSE MARKERS

(57) Abstract

The present invention relates to a composition comprising a plurality of nucleic acid molecules. The composition can be used as hybridizable array elements in a microarray. The present invention also relates to methods for screening compounds and therapeutics for metabolic responses indicative of a toxic compound.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TOXICOLOGICAL RESPONSE MARKERS

This application is filled under the Patent Cooperation Treaty and claims the benefits of U.S. Nonprovisional Application No. 09/141,825, our Docket No. PA-0010 US, filed on August 28, 1998, U.S. Nonprovisional Application No. 09/172,711, our Docket No. PA-0011 US, filed on October 13, 1998, andU.S. Nonprovisional Application No. 09/172,108, our Docket No. PA-0012 US, filed on October 13, 1998.

TECHNICAL FIELD

The present invention relates to compositions and methods for use in detecting metabolic and toxicological responses.

BACKGROUND ART

Toxicity testing is a mandatory and time-consuming part of the pharmaceutical drug development pipeline. A more rapid screen to determine the effects upon metabolism and to detect toxicity of lead drug candidates may be the use of gene expression microarrays. For example, microarrays of various kinds may be produced using full length genes or gene fragments. These arrays can then be used to test samples treated with the drug candidates to elucidate the gene expression pattern associated with drug treatment. This gene pattern can be compared with gene expression patterns associated with compounds which produce known toxicological and metabolic responses.

Benzo(a)pyrene is a known rodent and likely human carcinogen and is the prototype of a class of compounds, the polycyclic aromatic hydrocarbons (PAH). It is metabolized by several forms of cytochrome P450 (P450 isozymes) and associated enzymes to form both activated and detoxified metabolites. The ultimate metabolites are the bay-region diol epoxide, benzo(a)pyrene-7,8-diol-9,10-epoxide (BPDE) and the K-region diol epoxide, 9-hydroxy benzo(a)pyrene-4,5-oxide, both of which induce formation of DNA adducts. DNA adducts have been shown to persist in rat liver up to 56 days following treatment with benzo(a)pyrene at a dose of 10 mg/kg body weight three times per week for two weeks (Qu and Stacey (1996) Carcinogenesis 17:53-59).

Acetaminophen is a widely-used analgesic. It is metabolized by specific cytochrome P450 isozymes with the majority of the drug undergoing detoxification by glucuronic acid, sulfate and glutathione conjugation pathways. However, at supratherapeutic doses, acetaminophen is metabolized to an active intermediate, *N*-acetyl-*p*-benzoquinone imine (NAPQI) which can cause hepatic and renal failure. NAPQI then binds to sulfhydryl groups of proteins causing their inactivation and leading to subsequent cell death (Kroger et al. (1997) Gen. Pharmacol. 28:257-263).

Clofibrate is an hypolipidemic drug which lowers elevated levels of serum triglycerides.

In rodents, chronic treatment produces hepatomegaly and an increase in hepatic peroxisomes. Clofibrate has been shown to increase levels of cytochrome P450 4A and reduce the levels of P450 4F. It is also involved in transcription of β-oxidation genes as well as induction of peroxisome proliferator (PP) activated receptors (Kawashima et al. (1997) Arch. Biochem. Biophys. 347:148-154). Peroxisome proliferation that is induced by both clofibrate and the chemically-related compound fenofibrate is mediated by a common inhibitory effect on mitochondrial membrane depolarization (Zhou and Wallace (1999) Toxicol. Sci. 48:82-89).

The present invention provides compositions and methods for the screening of compounds for metabolic and toxicological responses.

DISCLOSURE OF INVENTION

The invention provides nucleic acid molecules whose transcript levels are modulated in a sample during a metabolic response to a toxic compound. The invention also provides nucleic acid molecules whose transcript levels are upregulated in a sample during a metabolic response to a toxic compound. The invention also provides nucleic acid molecules whose transcript levels are downregulated in a sample during a metabolic response to a toxic compound. Upregulation or downregulation is at least 2 fold, more preferably at least 2.5 fold, most preferably at least 3 fold. The metabolic response to a toxic compound may be a toxicological response.

In another aspect, the invention provides a method for screening a compound for a metabolic response to a test compound or molecule. The method comprises treating a tissue with a known toxic compound, determining levels of a plurality of nucleic acid molecules, selecting from the plurality of nucleic acid molecules those nucleic acid molecules that have levels modulated in samples treated with known toxic compounds when compared with untreated samples. Some of the transcript levels may be upregulated by a toxic compound, others may be downregulated by a toxic compound, and still others may be upregulated with one known toxic compound and be downregulated with another known toxic compound. The selected nucleic acid molecules which are upregulated and downregulated by a known toxic compound are arrayed upon a substrate. The method further comprises determining levels of nucleic acid molecules in a sample after the sample is treated with a compound. Levels of nucleic acid molecules in a sample so treated are then compared with the plurality of the arrayed nucleic acid molecules to identify which sample nucleic acid molecules are upregulated and downregulated by the compound.

Preferably, the comparing comprises contacting the arrayed nucleic acid molecules with the sample nucleic acid molecules under conditions effective to form hybridization complexes between the arrayed nucleic acid molecules and the sample nucleic acid molecules; and detecting the presence or absence of the hybridization complexes. In this context, similarity may mean that at least 1, preferably at least 5, more preferably at least 10, of the upregulated arrayed nucleic acid

molecules form hybridization complexes with the sample nucleic acid molecules at least once during a time course to a greater extent than would the probes derived from a sample not treated with the test compound or a known toxic compound. Similarity may also mean that at least 1, preferably at least 3, of the downregulated arrayed nucleic acid molecules form hybridization complexes with the sample nucleic acid molecules at least once during a time course to a lesser extent than would the sample nucleic acid molecules of a sample not treated with the test compound or a known toxic compound.

Preferred tissues are selected from the group consisting of liver, kidney, brain, spleen, pancreas and lung. Preferred toxic compounds are selected from the group consisting of hypolipidemic drugs, n-alkylcarboxylic acids, n-alkylcarboxylic acid precursors, azole antifungal compounds, leukotriene D4 antagonists, herbicides, pesticides, phthalate esters, phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol and their corresponding metabolites, acetaminophen and its corresponding metabolites, benzo(a)pyrene, 3-methylcholanthrene, benz(a)anthracene, 7,12-dimethylbenz(a)anthracene, their corresponding metabolites, and the like. The arrayed nucleic acid molecules comprise fragments of messenger RNA transcripts of genes that are up-regulated or down-regulated at least 2-fold, preferably at least 2.5-fold, more preferably at least 3-fold, in samples treated with known toxic compounds when compared with untreated samples. Preferred arrayed nucleic acid molecules are selected from the group consisting of SEQ ID NOs:1-117, or fragments thereof, some of whose levels are upregulated and others of whose levels are downregulated. More preferable are SEQ ID NOs: 3, 9, 10, 13, 19, 26, 31, 33, 35, 36, 37, 39, 42, 57, 67, 78, 81, 82, 94, and 98 which are upregulated, and SEQ ID NOs: 43, 49, 50, 52, 53, 54, 55, 56, 59, 61, 63, 68, 71, 74, 85, 87, 90, 95, 102, 103, 105, and 115 which are downregulated. Most preferable are SEQ ID NOs: 31, 33, 35, 36, 39, 52, 53, 54, 55, 63, 74, 81, 90, 94, and 95. In one embodiment, the polynucleotide targets are hybridizable array elements of a microarray.

Alternatively, the invention provides methods for screening a test compound or molecule for a metabolic response or for screening a sample for a metabolic response to a test compound or molecule.

Alternatively, the invention provides methods for screening a test compound or molecule for a previously unknown metabolic response or for screening a sample for a previously unknown metabolic response to a test compound or molecule.

In another aspect, the invention provides methods for preventing a toxicological response by administering complementary nucleotide molecules against one or more selected upregulated nucleic acid molecules or a ribozyme that specifically cleaves such molecules. Alternatively, a toxicological response may be prevented by administering sense nucleotide molecules for one or

more selected downregulated nucleic acid molecules.

In yet another aspect, the invention provides methods for preventing a toxicological response by administering an agonist which initiates transcription of a gene comprising a downregulated nucleic acid molecule of the invention. Alternatively, a toxicological response may be prevented by administering an antagonist which prevents transcription of a gene comprising an upregulated nucleic acid molecule of the invention.

The invention also provides a substantially purified mammalian protein or a portion thereof. The invention further provides isolated and purified proteins encoded by the nucleic acid molecules of SEQ ID NOs:1-117. Additionally, the invention provides a pharmaceutical composition comprising a substantially purified mammalian protein or a portion thereof in conjunction with a pharmaceutical carrier.

The invention further provides a method for using at least a portion of the proteins encoded by SEQ ID NOs:1-117 to produce antibodies. The invention also provides a method for using a protein or a portion thereof to screen a library of molecules to identify at least one ligand which specifically binds the protein, the method comprising combining the protein with the library of molecules under conditions allowing specific binding, and detecting specific binding, thereby identifying a ligand which specifically binds the protein. Such libraries include DNA and RNA molecules, peptides, agonists, antagonists, antibodies, immunoglobulins, drug compounds, pharmaceutical agents, and other ligands. In one aspect, the ligand identified using the method modulates the activity of the mammalian protein. In an analogous method, the protein or a portion thereof is used to purify a ligand. The method involves combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and separating the protein from the ligand to obtain purified ligand.

The invention further provides a method for inserting a marker gene into the genomic DNA of an animal to disrupt the expression of the natural nucleic acid molecule. The invention also provides a method for using the nucleic acid molecule to produce an animal model system, the method comprising constructing a vector containing the nucleic acid molecule; introducing the vector into a totipotent embryonic stem cell; selecting an embryonic stem cell with the vector integrated into genomic DNA; microinjecting the selected cell into a blastocyst, thereby forming a chimeric blastocyst; transferring the chimeric blastocyst into a pseudopregnant dam, wherein the dam gives birth to a chimeric animal containing at least one additional copy of nucleic acid molecule in its germ line; and breeding the chimeric animal to generate a homozygous animal model system.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

The Sequence Listing contains the nucleic acid sequence of exemplary nucleic acid molecules of the invention, SEQ ID NOs:1-117.

MODES FOR CARRYING OUT THE INVENTION

Definitions

"Sample" is used in its broadest sense. A sample containing nucleic acid molecules may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; a cell; a biological tissue or isolated fragment thereof, for example, a needle biopsy; a fingerprint or tissue print; natural or synthetic fibres; in a solution; in a liquid suspension; in a gaseous suspension; in an aerosol; and the like.

"Plurality" refers preferably to a group of one or more members, preferably to a group of at least about 10, and more preferably to a group of at least about 100 members, and even more preferably a group of 10,000 members.

"Substrate" refers to a rigid or semi-rigid support to which nucleic acid molecules or proteins are bound and includes membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, capillaries or other tubing, plates, polymers, and microparticles with a variety of surface forms including wells, trenches, pins, channels and pores.

"Modulates" refers to a change in activity (biological, chemical, or immunological) or lifespan resulting from specific binding between a molecule and either a nucleic acid molecule or a protein

"Microarray" refers to an ordered arrangement of hybridizable array elements on a substrate. The array elements are arranged so that there are preferably at least ten or more different array elements, more preferably at least 100 array elements, even more preferably at least 1000 array elements, and most preferably 10,000. Furthermore, the hybridization signal from each of the array elements is individually distinguishable. In a preferred embodiment, the array elements comprise nucleic acid molecules.

"Nucleic acid molecule" refers to a nucleic acid, oligonucleotide, nucleotide, polynucleotide or any fragment thereof. It may be DNA or RNA of genomic or synthetic origin, double-stranded or single-stranded, and combined with carbohydrate, lipids, protein or other

materials to perform a particular activity such as transformation or form a useful composition such as a peptide nucleic acid (PNA). "Oligonucleotide" is substantially equivalent to the terms amplimer, primer, oligomer, element, target, and probe and is preferably single stranded.

"Protein" refers to an amino acid sequence, oligopeptide, peptide, polypeptide or portions thereof whether naturally occurring or synthetic.

"Up-regulated" refers to a nucleic acid molecule whose levels increased in a treated sample compared with the nucleic acid molecule in an untreated sample.

"Down-regulated" refers to nucleic acid molecule whose levels decreased in a treated sample compared with the nucleic acid molecule in an untreated sample.

"Toxic compound" or "toxic agent" is any compound, molecule, or agent that elicits a biochemical, metabolic, and physiological response in an individual or animal, such as i) DNA damage, ii) cell damage, iii) organ damage or cell death, or iv) clinical morbidity or mortality .

"Toxicological response" refers to a biochemical, metabolic, and physiological response in an individual, animal, or test system which has been exposed to a toxic compound or toxic agent.

"Fragment" refers to an Incyte clone or any part of a nucleic acid molecule which retains a usable, functional characteristic. Useful fragments include oligonucleotides and polynucleotides which may be used in hybridization or amplification technologies or in regulation of replication, transcription or translation. Exemplary fragments are the first twenty consecutive nucleotides of SEQ ID NOs:1-117.

"Hybridization complex" refers to a complex between two nucleic acid molecules by virtue of the formation of hydrogen bonds between purines and pyrimidines.

"Ligand" refers to any molecule, agent, or compound which will bind specifically to a complementary site on a nucleic acid molecule or protein. Such ligands stabilize or modulate the activity of nucleic acid molecules or proteins of the invention and may be composed of at least one of the following: inorganic and organic substances including nucleic acids, proteins, carbohydrates, fats, and lipids.

"Substantially purified" refers to nucleic acid molecules or proteins that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free, from other components with which they are naturally associated.

The Invention

The present invention provides a composition and method of using the composition for screening test compounds and molecules for toxicological responses. Additionally the invention provides methods for characterizing the toxicological responses of a sample to a test compound or molecule. In particular, the present invention provides a composition comprising a plurality of

nucleic acid molecules derived from human cDNA libraries, monkey cDNA libraries, mouse cDNA libraries, normal rat liver cDNA libraries, normalized rat liver cDNA libraries and prehybridized rat liver cDNA libraries and rat kidney cDNA libraries in a test system. The nucleic acid molecules have been further selected for exhibiting up-regulated or down-regulated gene expression in rat livers when the rats have been exposed to a known hepatotoxin, including a peroxisome proliferator (PP), acetaminophen or one of its corresponding metabolites, and a polycyclic aromatic hydrocarbon (PAH).

PPs include hypolipidemic drugs, such as clofibrate, fenofibrate, clofenic acid, nafenopin, gemfibrozil, ciprofibrate, bezafibrate, halofenate, simfibrate, benzofibrate, etofibrate, WY-14,643, and the like; n-alkylcarboxylic acids, such as trichloroacetic acid, valproic acid, hexanoic acid, and the like; n-alkylcarboxylic acid precursors, such as trichloroethylene, etrachloroethylene, and the like; azole antifungal compounds, such as bifenazole, and the like; leukotriene D4 antagonists; herbicides; pesticides; phthalate esters, such as di-[2-ethylhexyl] phthalate, mono-[2-ethylhexyl] phthalate, and the like; and natural chemicals, such as phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol, and the like. In a prefered embodiment the toxic compound is clofibrate, or one of its corresponding metabolites. In another prefered embodiment the toxic compound is fenofibrate, or one of its corresponding metabolites.

PAHs include compounds such as benzo(a)pyrene, 3-methylcholanthrene, benzo(a)anthracene, 7,12-dimethylbenz(a)anthracene, their corresponding metabolites, and the like. In a preferred embodiment the toxic compound is benzo(a)pyrene, or one of its corresponding metabolites.

SEQ ID NOs:1-117 were identified by their pattern of at least two-fold up-regulation or down-regulation following hybridization with sample nucleic acid molecules from treated rat liver tissue. These and other nucleic acid molecules can be immobilized on a substrate as hybridizable array elements in a microarray format. The microarray may be used to characterize gene expression patterns associated with novel compounds to elucidate any metabolic responses or to monitor the effects of treatments during clinical therapy where metabolic responses to toxic compounds may be expected.

When the nucleic acid molecules are employed as hybridizable array elements in a microarray, the array elements are organized in an ordered fashion so that each element is present at a specified location on the substrate. Because the array elements are at specified locations on the substrate, the hybridization patterns and intensities (which together create a unique expression profile) can be interpreted in terms of expression levels of particular genes and can be correlated with a toxicological response associated with a test compound or molecule.

Furthermore, the present invention provides methods for screening test compounds and/or

molecules for potential toxicological responses and for screening a sample's toxicological response to a particular test compound or molecule. Briefly, these methods entail treating a sample with the test compound or molecule to elicit a change in gene expression patterns comprising the expression of a plurality of sample nucleic acid molecules. Nucleic acid molecules are selected by identifying those levels of expressed nucleic acid molecules in rat liver or kidney which are up-regulated or down-regulated at least 2-fold, more preferably at least 2.5-fold, most preferably at least 3-fold, when treated with a known toxic compound. The nucleic acid molecules are arrayed on a substrate. Then, the arrayed nucleic acid molecules and sample nucleic acid molecules are combined under conditions effective to form hybridization complexes which may be detected by methods well known in the art. Detection of higher or lower levels of such hybridization complexes compared with hybridization complexes derived from samples treated with a compound that is known not to induce a toxicological response correlates with a toxicological response to a test compound or a toxicological response to a molecule.

Complementary DNA libraries

Molecules are identified that reflect all or most of the genes that are expressed in rat tissue. Molecules may be identified by isolating clones derived from several types of rat cDNA libraries, including normal rat cDNA libraries, normalized rat cDNA libraries and prehybridized rat cDNA libraries. Clone inserts derived from these clones may be partially sequenced to generate expressed sequence tags (ESTs).

In one embodiment, two collections of ESTs are identified and sequenced. A first collection of ESTs (the originator molecules) are derived from rat liver and kidney and from the cDNA libraries presented in the Examples. A second collection includes ESTs derived from other rat cDNA libraries available in the ZOOSEQ database (Incyte Pharmaceuticals, Inc., Palo Alto CA).

The two collections of ESTs are clustered electronically to form master clusters of ESTs. Master clusters are formed by identifying overlapping EST molecules and assembling these ESTs. A nucleic acid fragment assembly tool, such as the Phrap tool (Phil Green, University of Washington) and the GELVIEW fragment assembly system (GCG, Madison WI), can be used for this purpose. The minimum number of clones which constitute a cluster is two. In another embodiment, a collection of human genes known to be expressed in response to toxic agents are used to select representative ESTs from the 113 rat cDNA libraries. The master cluster process is repeated for these molecules.

After assembling the clustered consensus nucleic acid sequences, a representative 5' clone is nominated from each master cluster. The most 5' clone is preferred because it is most likely to contain the complete gene. The nomination process is described in greater detail in "Relational

Database and System for Storing Information Relating to Biomolecular Sequences and Reagents", USSN 09/034,807, filed March 4, 1998, herein incorporated in its entirety by reference. The EST molecules are used as array elements on a microarray.

Selection of arrayed nucleic acid molecules

Samples are treated, preferably at subchronic doses, with one or more known toxic compounds over a defined time course. Preferably, the agents are peroxisomal proliferators (PPs), acetaminophen or one of its corresponding metabolites, and polycyclic aromatic hydrocarbons (PAHs).

The gene expression patterns derived from such treated biological samples can be compared with the gene expression patterns derived from untreated biological samples to identify nucleic acid molecules whose expression is either up-regulated or down-regulated due to the response to the toxic compounds. These molecules may then be employed as array elements alone or in combination with other array element molecules. Such a microarray is particularly useful to detect and characterize gene expression patterns associated with known toxic compounds. Such gene expression patterns can then be used for comparison to identify other compounds which also elicit a metabolic response to a toxic compound.

The arrayed nucleic acid molecules can be manipulated to optimize their performance in hybridization. To optimize hybridization, the arrayed nucleic acid molecules are examined using a computer algorithm to identify portions of genes without potential secondary structure. Such computer algorithms are well known in the art and are part of OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or LASERGENE software (DNASTAR, Madison WI). These programs can search within nucleic acid molecule sequences to identify stem loop structures and tandem repeats and to analyze G + C content of the sequence (those molecules with a G + C content greater than 60% are excluded). Alternatively, the arrayed nucleic acid molecules can be optimized by trial and error. Experiments can be performed to determine whether sample nucleic acid molecules and complementary arrayed nucleic acid molecules hybridize optimally under experimental conditions.

The arrayed nucleic acid molecules can be any RNA-like or DNA-like material, such as mRNAs, cDNAs, genomic DNA, peptide nucleic acids, branched DNAs and the like. The arrayed nucleic acid molecules can be in sense or antisense orientations.

In one embodiment, the arrayed nucleic acid molecules are cDNAs. The size of the DNA sequence of interest may vary, and is preferably from 50 to 10,000 nucleotides, more preferably from 150 to 3,500 nucleotides. In a second embodiment, the nucleic acid molecules are vector DNAs. In this case the size of the DNA sequence of interest, i.e., the insert sequence, may vary from about 50 to 10,000 nucleotides, more preferably from about 150 to 3,500 nucleotides.

The nucleic acid molecule sequences of the Sequence Listing have been prepared by current, state-of-the-art, automated methods and, as such, may contain occasional sequencing errors and unidentified nucleotides. Nucleotide analogues can be incorporated into the nucleic acid molecules by methods well known in the art. The only requirement is that the incorporated nucleotide analogues must serve to base pair with sample nucleic acid molecules. For example, certain guanine nucleotides can be substituted with hypoxanthine which base pairs with cytosine residues. However, these base pairs are less stable than those between guanine and cytosine. Alternatively, adenine nucleotides can be substituted with 2,6-diaminopurine which can form stronger base pairs than those between adenine and thymidine. Additionally, the nucleic acid molecules can include nucleotides that have been derivatized chemically or enzymatically. Typical modifications include derivatization with acyl, alkyl, aryl or amino groups.

The nucleic acid molecules can be immobilized on a substrate via chemical bonding. Furthermore, the molecules do not have to be directly bound to the substrate, but rather can be bound to the substrate through a linker group. The linker groups are typically about 6 to 50 atoms long to provide exposure to the bound nucleic acid molecule. Preferred linker groups include ethylene glycol oligomers, diamines, diacids and the like. Reactive groups on the substrate surface react with one of the terminal portions of the linker to bind the linker to the substrate. The other terminal portion of the linker is then functionalized for binding the nucleic acid molecule. Preferred substrates are any suitable rigid or semirigid support, including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which the arrayed nucleic acid molecules are bound.

The samples can be any sample comprising sample nucleic acid molecules and obtained from any bodily fluid (blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. The samples can be derived from any species, but preferably from eukaryotic species, and more preferably from mammalian species such as rat and human.

DNA or RNA can be isolated from the sample according to any of a number of methods well known to those of skill in the art. For example, methods of purification of nucleic acids are described in Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation, P. Tijssen, ed. Elsevier (1993). In one preferred embodiment, total RNA is isolated using the TRIZOL total RNA isolation reagent (Life Technologies, Gaithersburg MD) and mRNA is isolated using oligo d(T) column chromatography or glass beads. When sample nucleic acid molecules are amplified it is desirable to amplify the sample nucleic acid molecules and maintain the relative abundances of the original sample, including low abundance transcripts. RNA can be amplified in vitro, in situ or in vivo.

(See Eberwine US Patent No. 5,514,545).

It is also advantageous to include controls within the sample to assure that amplification and labeling procedures do not change the true distribution of nucleic acid molecules in the sample. For this purpose, a sample is spiked with an amount of control nucleic acid molecules predetermined to be detectable upon hybridization to its complementary arrayed nucleic acid molecule and the composition of nucleic acid molecules includes reference nucleic acid molecules which specifically hybridize with the control arrayed nucleic acid molecules. After hybridization and processing, the hybridization signals obtained should reflect accurately the amounts of control arrayed nucleic acid molecules added to the sample.

Prior to hybridization, it may be desirable to fragment the sample nucleic acid molecules. Fragmentation improves hybridization by minimizing secondary structure and cross-hybridization to other sample nucleic acid molecules in the sample or noncomplementary nucleic acid molecules. Fragmentation can be performed by mechanical or chemical means.

Labeling

The sample nucleic acid molecules may be labeled with one or more labeling moieties to allow for detection of hybridized arrayed/sample nucleic acid molecule complexes. The labeling moieties can include compositions that can be detected by spectroscopic, photochemical, biochemical, bioelectronic, immunochemical, electrical, optical or chemical means. The labeling moieties include radioisotopes, such as ^{32}P, ^{33}P or ^{35}S, chemiluminescent compounds, labeled binding proteins, heavy metal atoms, spectroscopic markers, such as fluorescent markers and dyes, magnetic labels, linked enzymes, mass spectrometry tags, spin labels, electron transfer donors and acceptors, and the like. Preferred fluorescent markers include Cy3 and Cy5 fluorophores (Amersham Pharmacia Biotech, Piscataway NJ).

Hybridization

The nulceic acid molecule sequence of SEQ ID NOs:1-117 and fragments thereof can be used in various hybridization technologies for various purposes in a test system. Hybridization probes may be designed or derived from SEQ ID NOs:1-117. Such probes may be made from a highly specific region such as the 5' regulatory region or from a conserved motif, and used in protocols to identify naturally occurring sequences encoding the mammalian protein, allelic variants, or related sequences, and should preferably have at least 50% sequence identity to any of the protein sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NOs:1-117 or from genomic sequences including promoters, enhancers, and introns of the mammalian gene. Hybridization or PCR probes may be produced using oligolabeling, nick translation, end-labeling, or PCR amplification in the presence of the labeled nucleotide. A vector containing the nucleic acid sequence may be used to

produce an mRNA probe in vitro by addition of an RNA polymerase and labeled nucleic acid molecules. These procedures may be conducted using commercially available kits such as those provided by Amersham Pharmacia Biotech.

The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. In solutions used for some membrane based hybridizations, addition of an organic solvent such as formamide allows the reaction to occur at a lower temperature. Hybridization can be performed at low stringency with buffers, such as 5 x SSC with 1% sodium dodecyl sulfate (SDS) at 60°C, which permits the formation of a hybridization complex between nucleotide sequences that contain some mismatches. Subsequent washes are performed at higher stringency with buffers such as 0.2 x SSC with 0.1% SDS at either 45°C (medium stringency) or 68°C (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acid sequences are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide can be added to the hybridization solution to reduce the temperature at which hybridization is performed, and background signals can be reduced by the use of other detergents such as Sarkosyl or Triton X-100 and a blocking agent such as salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (supra) and Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.

Hybridization specificity can be evaluated by comparing the hybridization of specificity-control nucleic acid molecules to specificity-control sample nucleic acid molecules that are added to a sample in a known amount. The specificity-control arrayed nucleic acid molecules may have one or more sequence mismatches compared with the corresponding arrayed nucleic acid molecules. In this manner, whether only complementary arrayed nucleic acid molecules are hybridizing to the sample nucleic acid molecules or whether mismatched hybrid duplexes are forming is determined.

Hybridization reactions can be performed in absolute or differential hybridization formats. In the absolute hybridization format, nucleic acid molecules from one sample are hybridized to the molecules in a microarray format and the signals detected after hybridization complex formation correlate to nucleic acid molecule levels in a sample. In the differential hybridization format, the differential expression of a set of genes in two biological samples is analyzed. For differential hybridization, nucleic acid molecules from both biological samples are prepared and labeled with different labeling moieties. A mixture of the two labeled nucleic acid molecules is added to a microarray. The microarray is then examined under conditions in which the emissions from the

two different labels are individually detectable. Molecules in the microarray that are hybridized to substantially equal numbers of nucleic acid molecules derived from both biological samples give a distinct combined fluorescence (Shalon et al. PCT publication WO95/35505). In a preferred embodiment, the labels are fluorescent markers with distinguishable excitation and emission spectra, such as Cy3 and Cy5 fluorophores.

After hybridization, the microarray is washed to remove nonhybridized nucleic acid molecules, then complex formation between the hybridizable array elements and the nucleic acid molecules is detected. Methods for detecting complex formation are well known to those skilled in the art. In a preferred embodiment, the nucleic acid molecules are labeled with a fluorescent label and measurement of levels and patterns of fluorescence indicative of complex formation is accomplished by fluorescence microscopy, preferably confocal fluorescence microscopy.

In a differential hybridization experiment, nucleic acid molecules from two or more different biological samples are labeled with two or more different fluorescent labels with different excitation and emission wavelengths. The labeled sample is excited with a specific excitation wavelength. Fluorescent signals are detected separately with different photomultipliers set to detect specific emission wavelengths. The relative abundances/expression levels of the nucleic acid molecules in two or more samples is obtained.

Typically, microarray fluorescence intensities can be normalized to take into account variations in hybridization intensities when more than one microarray is used under similar test conditions. In a preferred embodiment, individual arrayed-sample nucleic acid molecule complex hybridization intensities are normalized using the intensities derived from internal normalization controls contained on each microarray.

The labeled sample emits specific wavelengths which are detected using a plurality of photomultipliers. The relative abundances/expression levels of the arrayed nucleic acid molecules molecules can be used as hybridizable elements in a microarray. Such a microarray can be employed to identify expression profiles associated with particular toxicological responses. Then, a particular subset of these photomultipliers are set to detect specific wavelengths. The relative expression levels of the arrayed nucleic acid molecules can be identified as to which arrayed nucleic acid molecule expression is modulated in response to a particular toxicological agent. These photomultipliers set to detect specific wavelengths. The relative expression levels of the nucleic acid molecules can be employed to identify other compounds with a similar toxicological response.

Alternatively, for some treatments with known side effects, the microarray, and expression patterns derived therefrom, is employed to "fine tune" the treatment regimen. A dosage is established that minimizes expression patterns associated with undesirable side effects. This

approach may be more sensitive and rapid than waiting for the patient to show toxicological side effects before altering the course of treatment.

Generally, the method for screening a library of test compounds or molecules to identify those with a toxicological response entails selecting a plurality of arrayed nucleic acid molecule genes whose expression levels are modulated in tissues treated with known toxic compounds when compared with untreated tissues. Then a sample is treated with the test compound or molecule to induce a pattern of gene expression comprising the expression of a plurality of nucleic acid molecules. A test compound may be screened at several doses to determine which doses may be toxic and which may not.

Then, the expression levels of the arrayed nucleic acid molecules and the sample nucleic acid molecules are compared to identify those compounds that induce expression levels of the sample nucleic acid molecules that are similar to those of the arrayed nucleic acid molecules. In one preferred embodiment, gene expression levels are compared by contacting the arrayed nucleic acid molecules with the sample nucleic acid molecules under conditions effective to form hybridization complexes between arrayed nucleic acid molecules and sample nucleic acid molecules, and detecting the presence or absence of the hybridization complexes.

Similarity may mean that at least 1, preferably at least 5, more preferably at least 10, of the upregulated arrayed nucleic acid molecules form hybridization complexes with the sample nucleic acid molecules at least once during a time course to a greater extent than would the nucleic acid molecules of a sample not treated with the test compound. Similarity may also mean that at least 1, preferably at least 3, of the downregulated nucleic acid molecules form hybridization complexes with the nucleic acid molecules at least once during a time course to a lesser extent than would the nucleic acid molecules of a sample not treated with the test compound.

Such a similarity of expression patterns means that a toxicological response is associated with the test compound or molecule tested. Preferably, the toxic compounds belong to the class of peroxisomal proliferators (PPs), including hypolipidemic drugs, such as clofibrate, fenofibrate, clofenic acid, nafenopin, gemfibrozil, ciprofibrate, bezafibrate, halofenate, simfibrate, benzofibrate, etofibrate, WY-14,643, and the like; n-alkylcarboxylic acids, such as trichloroacetic acid, valproic acid, hexanoic acid, and the like; n-alkylcarboxylic acid precursors, such as trichloroethylene, etrachloroethylene, and the like; azole antifungal compounds, such as bifenazole, and the like; leukotriene D4 antagonists; herbicides; pesticides; phthalate esters, such as di-[2-ethylhexyl] phthalate, mono-[2-ethylhexyl] phthalate, and the like; and natural chemicals, such as phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol, and the like. In another embodiment, the toxic compounds are acetaminophen or one of its corresponding metabolites. In yet another embodiment, the toxic compound is a polycyclic aromatic hydrocarbon

(PAH), including compounds such as benzo(a)pyrene, 3-methylcholanthrene, benz(a)anthracene, 7,12-dimethylbenz(a)anthracene, their corresponding metabolites, and the like. Of particular interest is the study of the metabolic responses of these compounds on the liver, kidney, brain, spleen, pancreas, and lung.

Modification of Gene Expression Using Nucleic Acids

Gene expression may be modified by designing complementary or antisense molecules (DNA, RNA, or PNA) to the control, 5', 3', or other regulatory regions of the mammalian gene. Oligonucleotides designed with reference to the transcription initiation site are preferred. Similarly, inhibition can be achieved using triple helix base-pairing, which inhibits the binding of polymerases, transcription factors, or regulatory molecules (Gee et al. In: Huber and Carr (1994) Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary molecule may also be designed to block translation by preventing binding between ribosomes and mRNA. In one alternative, a library of nucleic acid molecules or fragments thereof may be screened to identify those which specifically bind a regulatory, nontranslated sequence.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA followed by endonucleolytic cleavage at sites such as GUA, GUU, and GUC. Once such sites are identified, an oligonucleotide with the same sequence may be evaluated for secondary structural features which would render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing their hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary nucleic acids and ribozymes of the invention may be prepared via recombinant expression, in vitro or in vivo, or using solid phase phosphoramidite chemical synthesis. In addition, RNA molecules may be modified to increase intracellular stability and half-life by addition of flanking sequences at the 5' and/or 3' ends of the molecule or by the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Modification is inherent in the production of PNAs and can be extended to other nucleic acid molecules. The inclusion of nontraditional bases such as inosine, queosine, and wybutosine, or the modification of adenine, cytidine, guanine, thymine, and uridine with acetyl-, methyl-, thio- groups, renders the molecule less available as a substrate to endogenous endonucleases.

Screening Assays

The nucleic acid molecule encoding the mammalian protein may be used to screen a library of molecules for specific binding affinity. The libraries may be DNA molecules, RNA molecules, PNAs, peptides, proteins such as transcription factors, enhancers, repressors, and other

ligands which regulate the activity, replication, transcription, or translation of the nucleic acid molecule in the biological system. The assay involves combining the mammalian nucleic acid molecule or a fragment thereof with the library of molecules under conditions allowing specific binding, and detecting specific binding to identify at least one molecule which specifically binds the nucleic acid molecule.

Similarly the mammalian protein or a portion thereof may be used to screen libraries of molecules in any of a variety of screening assays. The portion of the protein employed in such screening may be free in solution, affixed to an abiotic or biotic substrate, or located intracellularly. Specific binding between the protein and molecule may be measured. Depending on the kind of library being screened, the assay may be used to identify DNA, RNA, or PNA molecules, agonists, antagonists, antibodies, immunoglobulins, inhibitors, peptides, proteins, drugs, or any other ligand, which specifically binds the protein. One method for high throughput screening using very small assay volumes and very small amounts of test compound is described in USPN 5,876,946, incorporated herein by reference, which screens large numbers of molecules for enzyme inhibition or receptor binding.

Purification of Ligand

The nucleic acid molecule or a fragment thereof may be used to purify a ligand from a sample. A method for using a mammalian nucleic acid molecule or a fragment thereof to purify a ligand would involve combining the nucleic acid molecule or a fragment thereof with a sample under conditions to allow specific binding, detecting specific binding, recovering the bound protein, and using an appropriate agent to separate the nucleic acid molecule from the purified ligand.

Similarly, the protein or a portion thereof may be used to purify a ligand from a sample. A method for using a mammalian protein or a portion thereof to purify a ligand would involve combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and using an appropriate chaotropic agent to separate the protein from the purified ligand.

Pharmacology

Pharmaceutical compositions are those substances wherein the active ingredients are contained in an effective amount to achieve a desired and intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compound, the therapeutically effective dose may be estimated initially either in cell culture assays or in animal models. The animal model is also used to achieve a desirable concentration range and route of administration. Such information may then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of protein or inhibitor which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity of such agents may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it may be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indexes are preferred. The data obtained from cell culture assays and animal studies are used in formulating a range of dosage for human use.

Model Systems

Animal models may be used as bioassays where they exhibit a toxic response similar to that of humans and where exposure conditions are relevant to human exposures. Mammals are the most common models, and most toxicity studies are performed on rodents such as rats or mice because of low cost, availability, and abundant reference toxicology. Inbred rodent strains provide a convenient model for investigation of the physiological consequences of under- or over-expression of genes of interest and for the development of methods for diagnosis and treatment of diseases. A mammal inbred to over-express a particular gene, so that the protein is secreted in milk, may also serve as a convenient source of the protein expressed by that gene.

Toxicology

Toxicology is the study of the effects of test compounds, molecules, or toxic agents on living systems to identify adverse effects. The majority of toxicity studies are performed on rats or mice to help predict whether adverse effects of agents will occur in humans. Observation of qualitative and quantitative changes in physiology, behavior, homeostatic, developmental, and reproductive processes, and lethality are used to generate profiles of safe or toxic responses and to assess the consequences on human health following exposure to the agent.

Genetic toxicology identifies and analyzes the ability of an agent to produce damage at a cellular or subcellular level. Such genotoxic agents usually have common chemical or physical properties that facilitate interaction with nucleic acids and are most harmful when mutated chromosomes are passed along to progeny. Toxicological studies may identify agents that increase the frequency of structural or functional abnormalities in progeny if administered to either parent before conception, to the mother during pregnancy, or to the developing organism. Mice and rats are most frequently used in these tests because of their short reproductive cycle which produces the number of organisms needed to satisfy statistical requirements.

Acute toxicity tests are based on a single administration of the agent to the subject to determine the symptomology or lethality of the agent. Three experiments are conducted: 1) an initial dose-range-finding experiment, 2) an experiment to narrow the range of effective doses,

and 3) a final experiment for establishing the dose-response curve.

Prolonged toxicity tests are based on the repeated administration of the agent. Rat and dog are commonly used in these studies to provide data from species in different taxonomic orders. With the exception of carcinogenesis, there is considerable evidence that daily administration of an agent at high-dose concentrations for periods of three to four months will reveal most forms of toxicity in adult animals. Chronic toxicity tests, with a duration of a year or more, are used to demonstrate either the absence of toxicity or the carcinogenic potential of an agent. When studies are conducted on rats, a minimum of at least one test group plus one control group are used. Animals are quarantined, examined for health, and monitored at the outset and at intervals throughout the experiment.

Transgenic Animal Models

Transgenic rodents which overexpress or underexpress a gene of interest may be inbred and used to model human diseases or to test compounds and molecules for therapeutic or toxicological effects. (See USPN 4,736,866; USPN 5,175,383; and USPN 5,767,337; incorporated herein by reference). In some cases, the introduced gene may be activated at a specific time in a specific tissue type during fetal development or postnatally. Expression of the transgene is monitored by analysis of phenotype or tissue-specific mRNA expression, in transgenic animals before, during, and after being challenged with experimental drug therapies.

Embryonic Stem Cells

Embryonic stem cells (ES) isolated from rodent embryos retain the potential to form an embryo. When ES cells are placed inside a carrier embryo, they resume normal development and contribute to all tissues of the live-born animal. ES cells are the preferred cells used in the creation of experimental knockout and knockin rodent strains. Mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and are grown under culture conditions well known in the art. Vectors for knockout strains contain a disease gene candidate modified to include a marker gene which disrupts transcription and/or translation in vivo. The vector is introduced into ES cells by transformation methods such as electroporation, liposome delivery, microinjection, and the like which are well known in the art. The endogenous rodent gene is replaced by the disrupted disease gene through homologous recombination and integration during cell division. Then transformed ES cells are selected under conditions, identified, and preferably microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains.

ES cells are also used to study the differentiation of various cell types and tissues in vitro, such as neural cells, hematopoietic lineages, and cardiomyocytes (Bain et al. (1995) Dev. Biol.

168:342-357; Wiles and Keller (1991) Development 111:259-267; and Klug et al. (1996) J. Clin. Invest. 98:216-224). Recent developments demonstrate that ES cells derived from human blastocysts may also be manipulated in vitro to differentiate into eight separate cell lineages, including endoderm, mesoderm, and ectodermal cell types (Thomson (1998) Science 282:1145-1147).

Knockout Analysis

In gene knockout analysis, a region of a human disease gene candidate is enzymatically modified to include a non-mammalian gene such as the neomycin phosphotransferase gene (*neo*; Capecchi (1989) Science 244:1288-1292). The inserted coding sequence disrupts transcription and translation of the targeted gene and prevents biochemical synthesis of the disease candidate protein. The modified gene is transformed into cultured embryonic stem cells (described above), the transformed cells are injected into rodent blastulae, and the blastulae are implanted into pseudopregnant dams. Transgenic progeny are crossbred to obtain homozygous inbred lines.

Knockin Analysis

Totipotent ES cells, present in the early stages of embryonic development, can be used to create knockin humanized animals (pigs) or transgenic animal models (mice or rats) of human diseases. With knockin technology, a region of a human gene is injected into animal ES cells, and the human sequence integrates into the animal cell genome by recombination. Totipotent ES cells which contain the integrated human gene are handled as described above. Inbred animals are studied and treated to obtain information on the analogous human condition. These methods have been used to model several human diseases. (See, e.g., Lee et al. (1998) Proc. Natl. Acad. Sci. 95:11371-11376; Baudoin et al. (1998) Genes Dev. 12:1202-1216; and Zhuang et al. (1998) Mol. Cell Biol. 18:3340-3349).

Non-Human Primate Model

The field of animal testing deals with data and methodology from basic sciences such as physiology, genetics, chemistry, pharmacology and statistics. These data are paramount in evaluating the effects of test compounds or molecules on non-human primates as they can be related to human health. Monkeys are used as human surrogates in vaccine and drug evaluations, and their responses are relevant to human exposures under similar conditions. Cynomolgus and rhesus monkeys (Macaca fascicularis and Macaca mulatta, respectively) and common marmosets (Callithrix jacchus) are the most common non-human primates (NHPs) used in these investigations. Since great cost is associated with developing and maintaining a colony of NHPs, early research and toxicological studies are usually carried out in rodent models. In studies using behavioral measures such as drug addiction, NHPs are the first choice test animal. In addition, NHPs and individual humans exhibit differential sensitivities to many drugs and toxins and can be

classified as "extensive metabolizers" and "poor metabolizers" of these agents.

In additional embodiments, the nucleic acid molecules which encode the mammalian protein may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleic acid molecules that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Examples

It is understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary. It is also understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims. The examples below are provided to best describe the subject invention and its representative constituents.

I cDNA Library Construction

The RALINOT01 cDNA library was constructed from liver tissue removed from a pool of fifty 10- to 11-week-old Sprague-Dawley female rats (Pharmacon, Waverly PA). The animals were housed in standard laboratory caging and fed PMI-certified Rodent Diet #5002. The animals appeared to be in good health at the time tissue was harvested. The animals were anesthetized by CO_2 inhalation, and then cardiocentesis was performed.

Frozen tissue was homogenized and lysed in TRIZOL reagent (1 g tissue/10 ml TRIZOL; Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate, using a POLYTRON homogenizer (PT-3000; Brinkmann Instruments, Westbury NY). After a brief incubation on ice, chloroform (1:5 v/v) was mixed with the reagent, and then centrifuged at 1,000 rpm. The upper aqueous layer was removed to a fresh tube, and the RNA precipitated with isopropanol, resuspended in DEPC-treated water, and treated with DNase I for 25 min at 37°C. The RNA was re-extracted once with phenol-chloroform, pH 4.7, and precipitated using 0.3 M sodium acetate and 2.5 volumes ethanol. The mRNA was then isolated using an OLIGOTEX kit (QIAGEN, Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SUPERSCRIPT plasmid system (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL-4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into the pINCY1 plasmid vector (Incyte Pharmaceuticals). The plasmid pINCY1 was subsequently transformed into DH5α or DH10B competent cells (Life Technologies).

The RAKINOT01 library was constructed using mRNA isolated from kidney tissue removed from a pool of fifty, 7- to 8-week-old male Sprague-Dawley rats, as described above.

The RAKINOT02 library was constructed using mRNA isolated from kidney tissue removed from a pool of fifty, 10- to 11-week-old female Sprague-Dawley rats, as described above.

II cDNA Library N rmalizati n

In some cases, cDNA libraries were normalized in a single round according to the procedure of Soares et al. (1994, Proc. Natl. Acad. Sci. 91:9228-9232) with the following modifications. The primer to template ratio in the primer extension reaction was increased from 2:1 to 10:1. Reduction of each dNTP concentration in the reaction to 150μM allowed the generation of longer (400-1000 nucleotide (nt)) primer extension products. The reannealing hybridization was extended from 13 to 19 hours. The single stranded DNA circles of the normalized library were purified by hydroxyapatite chromatography, converted to partially double-stranded by random priming, and electroporated into DH10B competent bacteria (Life Technologies).

The Soares normalization procedure is designed to reduce the initial variation in individual cDNA frequencies and to achieve abundances within one order of magnitude while maintaining the overall sequence complexity of the library. In the normalization process, the prevalence of high-abundance cDNA clones decreases significantly, clones with mid-level abundance are relatively unaffected, and clones for rare transcripts are increased in abundance. In the modified Soares normalization procedure, significantly longer hybridization times are used to increase gene discovery rates by biasing the normalized libraries toward low-abundance cDNAs that are well represented in a standard transcript image.

The RALINON03, RALINON04, and RALINON07 normalized rat liver cDNA libraries were constructed with 2.0 x 10^6, 4.6 x 10^5, and 2.0 x 10^6 independent clones from the RALINOT01 cDNA library, respectively. The RALINOT01 cDNA library was normalized in one round using conditions adapted from Soares (supra) except that a significantly longer (48-hour) reannealing hybridization was used.

III cDNA Library Prehybridization

The RALINOH01 cDNA library was constructed with clones from the RALINOT01 cDNA library. After preparation of the RALINOT01 cDNA library, 9,984 clones were spotted onto a nylon filter, lysed, and the plasmid DNA was bound to the filter. The filter was incubated with pre-warmed hybridization buffer and then hybridized at 42°C for 14-16 hours in 0.75 M NaCl, 0.1 M Na_2HPO_4/NaH_2PO_4, 0.15 M tris-HCl (pH 7.5), 5x Denhardt's Solution, 2% SDS, 100 µg/ml sheared salmon sperm DNA, 50% formamide, and [^{32}P]-labeled oligonucleotide molecules made from reverse transcribed rat liver mRNA from an untreated animal. The filter was rinsed with 2 x SSC (saline sodium citrate) at ambient temperature for 5 minutes followed by washing for 30 minutes at 68°C with pre-warmed washing solution (2 x SSC, 1% SDS). The wash was

repeated with fresh washing solution for an additional 30 minutes at 68°C. Filters were then washed twice with pre-warmed washing solution (0.6 x SSC, 1% SDS) for 30 minutes at 68°C. Some 4,224 clones had very low hybridization signals and about 20% of the clones had no signals and two groups were isolated and sequenced.

IV Isolation and Sequencing of cDNA Clones

DNA was isolated using the following protocol. Single bacterial colonies were transferred into individual wells of 384-well plates (Genetix Ltd, Christchurch, United Kingdom) using sterile toothpicks. The wells contained 1 ml of sterile Terrific Broth (Life Technologies) with 25 mg/l carbenicillin and 0.4% glycerol (v/v). The plates were covered and placed in an incubator (Thermodyne, Newtown Square PA) at 37°C for 8-10 hours. Plasmid DNA was released from the cells and amplified using direct link PCR (Rao, V.B. (1994) Anal. Biochem. 216:1-14) as follows. The direct link PCR solution included 30 ml of NUCLEIX PLUS PCR nucleotide mix (Amersham Pharmacia Biotech, Piscataway NJ) and 300 µl of Taq DNA polymerase (Amersham Pharmacia Biotech). Five microlitres of the PCR solution were added to each of the 384 wells using the MICROLAB 2200 system (Hamilton, Reno NV); plates were centrifuged at 1000 rpm for 20 seconds and refrigerated until use. A 384 pin tool (V&P Scientific Inc, San Diego CA) was used to transfer bacterial cells from the incubation plate into the plate containing the PCR solution where 0.1% Tween 20 caused the cells to undergo lysis and release the plasmid DNA. After lysis, the plates were centrifuged up to 500 rpm, covered with a cycle sealer, and cycled using a 384-well DNA ENGINE thermal cycler (MJ Research, Watertown MA) using the program dPCR30 with the following parameters: Step 1) 95°C, 1 minute; Step 2) 94°C, 30 seconds; Step 3) 55°C, 30 seconds; Step 4) 72°C, 2 minutes; Step 5) steps 2, 3, and 4 repeated 29 times; Step 6) 72°C, 10 minutes; and Step 7) storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICO GREEN quantitation reagent (0.25% (v/v), Molecular Probes, Eugene OR) dissolved in 1x TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the quantitation reagent. The plate was scanned in a Fluoroscan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantitate the concentration of DNA. Typical concentrations of each DNA sample were in the range of 100 to 500 ng/ml.

The cDNAs were prepared for sequencing using either a HYDRA microdispenser (Robbins Scientific, Sunnyvale CA) or MICROLAB 2200 system (Hamilton) in combination with the DNA ENGINE thermal cyclers (MJ Research). The cDNAs were sequenced using the method of Sanger, F. and A.R. Coulson (J. Mol. Biol. (1975) 94:441-448) and the ABI 377 sequencing systems (PE Biosystems). Most of the isolates were sequenced according to standard ABI

protocols using ABI kits (PE Biosystems). The solution volumes were used at 0.25x - 1.0x concentrations. Typically, 500 to 700 base pairs were sequenced in 3.5 to 4 hours. In the alternative, cDNAs may have been sequenced using solutions and dyes from Amersham Pharmacia Biotech.

V Rat Liver and Kidney Gene Selection

As a first step, originator molecules from high throughput sequencing experiments were derived from clone inserts from RALINOT01, RAKINOT01, RAKINOT02, RALINOH01, RALINON03, RALINON04 and RALINON07. cDNA library clones were obtained. There were 18,140 rat liver molecules and 5,779 rat kidney molecules.

Additionally, 1,500 rat molecules derived from clone inserts of any of 113 rat cDNA libraries were selected based on their homology to genes coding for polypeptides implicated in toxicological responses including peroxisome-associated genes, lysosome-associated genes, apoptosis-associated genes, P450 cytochromes, detoxification genes such as sulfotransferases, glutathione S-transferase, and cysteine proteases, and the like.

Then, all the remaining molecules derived from all of the rat cDNA library clones were clustered based on the originator molecules described above. The clustering process involved identifying overlapping molecules that have a match quality indicated by a product score of 50 using BLAST.

6581 master clusters were identified.

After forming the clone clusters, a consensus sequence was generated based on the assembly of the clone molecules using Phrap (Phil Green, University of Washington). The assembled molecules were then annotated by first screening the assembled molecules against GenBank using BLASTn and then by screening the assembled molecules against GenPept using FASTX. About two thirds of the assembled molecules were annotated, about one third of the assembled molecules were not annotated.

VI Substrate and Array Element/Probe Preparation

Clones nominated in the process described in Example V were used to generate array elements. Each array element was amplified from bacterial cells. PCR amplification used primers complementary to the vector sequences flanking the cDNA insert. Array elements were amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements were then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements were immobilized on polymer-coated glass slides. Glass microscope slides (Corning, Corning NY) cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides were etched in 4%

hydrofluoric acid (VWR, West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma-Aldrich, St. Louis MO) in 95% ethanol. Coated slides were cured in a 110°C oven.

Array elements were applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522 and incorporated herein by reference. In brief, 1 µl of the array element DNA, at an average concentration of 0.5 µg/ml in 3 x SSC, was loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposited about 5 nl of the array element sample per slide. A total of 7404 array elements representing rat liver and kidney genes and a variety of control elements, including 14 synthetic control molecules, human genomic DNA, and yeast genomic DNA, were arrayed in four identical quadrants within a 1.8 cm^2 area of the glass substrate.

Microarrays were UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays were washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites were blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

VII Target Preparation

Male Sprague-Dawley rats (6-8 wk old) were dosed intraperitoneally with clofibrate (CLO; Acros, Geel, Belgium) at 250 mg/kg body weight (bw), acetaminophen (APAP; Acros) at 1000 mg/kg bw, benzo(a)pyrene (B(a)P; Acros) at 10 mg/kg bw, or dimethylsulfoxide vehicle (DMSO; Acros) at less than 2 ml/kg bw and the animals were later euthanized by CO_2 inhalation. Animals were monitored daily for physical condition and body weight. Three animals per group were sacrificed approximately 12 hours, 1 day (d), 3d, 7d, 14d, and 28d following the single dose. Prior to sacrifice a blood sample from each animal was drawn and assayed for serum alanine transferase (ALT) and aspartate aminotransferase (AST) levels using a diagnostic kit (Sigma-Aldrich). Observed gross pathology and liver weights were recorded at time of necropsy. Liver, kidney, brain, spleen and pancreas from each rat were harvested, flash frozen in liquid nitrogen, and stored at -80°C.

For each probe preparation, frozen liver was homogenized and lysed in TRIZOL reagent (Life Technologies, Gaithersburg MD) following the modifications for liver RNA isolation. Messenger RNA was isolated using an OLIGOTEX kit (QIAGEN) and labeled with either Cy3- or Cy5-labeled primers (Operon Technologies, Alameda CA) using the GEMBRIGHT labeling kit (Incyte Pharmaceuticals). Messenger RNA isolated from tissues of rats treated with clofibrate, acetaminophen, or benzo(a)pyrene was labeled with Cy5 and mRNA isolated from tissues of rats treated with DMSO was labeled with Cy3. Quantitative and differential expression pattern control

cDNAs were added to each labeling reaction. Labeled cDNA was treated with 0.5 M sodium bicarbonate (pH 9.2) for 20 min at 85°C to degrade the RNA and purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA). Cy3-labeled control sample and Cy5-labeled experimental sample were combined and precipitated in glycogen, sodium acetate, and ethanol.

Probes are also prepared from tissue needle biopsy samples. Samples are used to identify changes within the tissue following exposure to, for example, a toxic compond, a potential toxic compound, a compound with unknown metabolic responses, or a pharmacological compound.

VIII Hybridization

Hybridizations were carried out using the methods described by Shalon (supra).

IX Detection

The microscope used to detect the reporter-labeled hybridization complexes was equipped with an Innova 70 mixed gas 10 W laser (Coherent Lasers, Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3, and 632 nm for excitation of Cy5. The excitation laser light was focused on the array using a 20x microscope objective (Nikon, Melville NY). The slide containing the array was placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example was scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excited the two fluorophores sequentially. Emitted light was split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics, San Jose CA) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes were used to filter the signals. The emission maxima of the fluorophores used were 565 nm for Cy3 and 650 nm for Cy5. Each array was typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus was capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans was typically calibrated using the signal intensity generated by a cDNA control species added to the probe mix at a known concentration. A specific location on the array contained a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two probes from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration was done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube was digitized using a 12-bit RTI-835H analog-to-

digital (A/D) conversion board (Analog Devices, Norwood MA) installed in an IBM-compatible PC computer. The digitized data were displayed as an image where the signal intensity was mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data was also analyzed quantitatively. Where two different fluorophores were excited and measured simultaneously, the data were first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid was superimposed over the fluorescence signal image such that the signal from each spot was centered in each element of the grid. The fluorescence signal within each element was then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis was the GEMTOOLS gene expression analysis program (Incyte Pharmaceuticals). In one analysis, where two different samples were prepared from identically treated cell cultures, expression patterns of those cDNAs which changed between 1.6- and 1.7-fold were within the 95% confidence limits of a Poisson normal distribution profile (T. Theriault, pers. communication).

X Results

The expression patterns of eight cytochrome P450 isozymes known to be induced in a toxicological response were monitored during the 28 day time course. The results using clofibrate, acetaminophen, and benzo(a)pyrene are shown in Table 1, Table 2, and Table 3, respectively. Each of the known genes was upregulated greater than 2 fold at least once during the time course.

TABLE 1 Gene expression patterns (x fold increase) of known genes in clofibrate-treated rat liver

Gene	12 hours	1 day	3 days	7 days	28 days
P450 LA-omega	15	26	2.0	2.1	3.0
P450 4A	6.5	16.5	2.1	3.0	3.5
P450 3A	0.14	1.6	0.63	0.50	0.45

TABLE 2 Gene expression patterns (x fold increase) of known genes in acetaminophen-treated rat liver

Gene	12 hours	24 hours	3 days	7 days	14 days	28 days
P450A	1	4.4	2.2	2.0	4.6	4.8
P450F	0.50	0.23	2.0	1.8	2.2	2.2
P450 14DM	0.45	0.32	2.2	1.6	1.8	0.56

TABLE 3 Gene expression patterns (x fold increase) of known genes in benzo(a)pyrene-treated rat liver

Gene	12 hours	1 day	3 days	7 days	14 days	28 days
P450 LA-omega	1.2	2.3	2.4	1.4	6.8	1.2
P450 MCA-inducible	8.2	11.8	4.4	2.2	2.4	1.2
P450 ISF/B-NF	9.6	7.4	6.2	2.4	2.4	1.2

We have discovered novel nucleotide molecules that are up-regulated or down-regulated at least 2-fold at least once during the time course. These molecules are SEQ ID NOs:1-117 provided in the Sequence Listing. These polynucleotide molecules can be used for screening test compounds or molecules for a toxicologic effect.

Table 4 shows the gene expression pattern of selected molecules that were upregulated at least 2-fold at least once during the time course following treatment with clofibrate (CLO) and Table 5 shows the gene expression pattern of selected molecules that were downregulated at least 2-fold at least once during the time course following treatment with clofibrate.

TABLE 4 Gene expression patterns (x fold increase) of CLO-upregulated nucleic acid moleucles

SEQ ID NO:	12 hours	1 day	3 days	7 days	28 days
35	11.6	14.4	2.4	3.0	3.2
36	11.6	18.7	3.0	3.3	3.8
31	1.2	2.8	1.0	2.3	4.8
57	0.9	1.9	0.9	1.5	4.5
67	4.3	1.1	1.6	1.7	5.7
81	5.1	1.2	1.7	1.8	6.0
94	4.8	1.4	2.0	1.5	2.4
33	5.1	1.3	1.9	1.8	5.5

TABLE 5 Gene expression patterns (x fold increase) of CLO-downregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	28 days
102	0.15	1.4	1.0	0.77	0.67
103	0.13	1.2	0.83	0.63	0.56
52	0.13	0.56	0.37	1.0	1.2
43	0.13	1.1	0.91	0.71	0.56
53	0.11	0.67	0.36	1.0	1.2
54	0.14	0.63	0.59	1.1	0.29
55	0.16	0.67	0.71	1.2	0.32
63	0.33	0.14	1.1	0.83	1.2
105	0.14	1.2	1.0	0.77	0.71
68	0.16	0.67	0.53	1.1	1.4
71	0.43	0.18	0.40	0.34	0.23
74	0.06	0.71	0.42	1.1	1.2
115	0.22	1.5	0.77	1.7	1.3
85	0.19	0.45	1.0	1.3	1.8
90	0.12	0.48	1.2	1.0	1.2
95	0.14	0.91	0.56	1.5	1.4

Table 6 shows the gene expression pattern of selected molecules that were upregulated at least 2-fold at least once during the time course following treatment with acetaminophen (APAP) and Table 7 shows the gene expression pattern of selected molecules that were downregulated at least 2-fold at least once during the time course following treatment with acetaminophen.

TABLE 6 Gene expression patterns (x fold increase) of APAP-upregulated nucleic acid molecules

SEQ ID NO:	12 hours	24 hours	3 days	7 days	14 days	28 days
35	3.1	6.6	2.9	3.3	4.9	7.5
36	4.7	10.1	4.0	4.2	6.9	9.8
78	0.9	4.4	1.2	1.5	1.1	1.4
81	2.9	5.1	1.4	1.8	2.3	2.4
82	1.2	4.2	1.3	1.0	1.7	1.4
39	2.4	9.0	2.6	1.7	2.2	2.4
94	1.2	4.9	1.2	1.1	2.0	2.0
33	4.3	5.9	1.5	1.7	2.9	3.2
98	1.3	6.1	1.5	1.9	1.8	2.1

TABLE 7 Gene expression patterns (x fold increase) of APAP-downregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	14 days	28 days
49	0.59	0.15	1.2	1.0	0.83	1.1
50	0.83	0.37	0.43	0.37	0.22	0.2
52	0.63	0.08	1.0	0.71	0.83	0.45
53	0.25	0.07	1.1	0.71	0.83	0.42
54	0.43	0.19	0.04	0.71	0.29	0.36
55	0.35	0.22	0.07	0.77	0.31	0.42
56	0.38	0.21	0.5	0.32	1.1	1.1
59	0.18	0.77	2.5	1.4	1.2	1.6
61	0.15	0.53	0.91	0.71	0.71	1.8
63	0.13	0.05	0.23	0.77	0.43	0.77
74	0.19	0.09	1.1	1.0	1.4	0.56
87	0.42	0.10	0.53	0.63	0.63	0.67
90	0.16	0.29	1.2	0.77	0.83	1.1
95	0.22	0.20	2.7	1.7	1.6	1.0

Table 8 shows the gene expression pattern of selected molecules that were upregulated at least 2-fold at least once during the time course following treatment with benzo(a)pyrene (B(a)P) and Table 9 shows the gene expression pattern of selected molecules that were downregulated at least 2-fold at least once during the time course following treatment with benzo(a)pyrene.

TABLE 8 Gene expression patterns of B(a)P-upregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	14 days	28 days
3	3.4	1.9	0.7	0.5	1.99	0.77
9	1.6	3.2	1.2	1.1	3	1.5
10	2.8	5.9	3.2	2.1	2.9	1.8
13	2.9	6.1	3.1	2.3	3.3	1.9
19	2.7	3.5	3	1.9	1.7	1.5
26	1.1	4.7	1.5	1.3	5	2
31	2.3	3.8	1.6	2	1.7	2.1
33	2.1	4.1	3.2	2	1.7	1.6
35	1.2	3	5.1	1.4	5	1.3
37	3.4	0.5	0.6	0.7	0.9	0.5
39	1.5	3.5	1.8	1.5	3.5	2.1
42	9.1	9.1	5.2	2.4	2.1	1

TABLE 9 Gene expression patterns of B(a)P-downregulated nulciec acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	14 days	28 days
11	0.3	0.5	0.4	0.3	0.53	0.53
25	0.3	0.9	0.5	0.7	0.42	2.1
27	1	0.1	1	1.1	0.09	0.53
28	0.3	0.3	1.2	1.2	0.77	1.1
45	1.2	0.2	0.4	0.6	0.77	0.37

CLAIMS

What is claimed is:

1. A method for detecting or diagnosing the effect of a test compound or molecule associated with increased or decreased levels of a nucleic acid molecule in a mammalian subject comprising:

a) treating a mammalian subject with a toxic compound or molecule;

b) obtaining a sample containing nucleic acids from the mammalian subject treated with the toxic compound or molecule;

c) contacting the sample with a microarray comprising a plurality of nucleic acid molecules comprising SEQ ID NOs: 1-117, or a fragment thereof, under conditions for the formation of one or more hybridization complexes;

d) detecting the hybridization complexes, wherein the presence, absence or change in amount of the hybridization complex, as compared with the hybridization complexes formed from nucleic acid molecules from an untreated mammalian subject, is indicative of a metabolic response to the toxic compound or molecule;

e) measuring the level of nucleic acid molecules in a sample from a mammalian subject treated with a test compound or molecule using the method of steps (c) and (d); and

f) comparing the level detected in step (e) to a level of nucleic acid molecules present in normal or untreated biological sample in which an increase or decrease in the level of nucleic acid molecule as compared to normal levels indicates a toxicological response.

2. The method of claim 1 wherein the toxic compound or molecule is selected from hypolipidemic drugs, n-alkylcarboxylic acids, n-alkylcarboxylic acid precursors, azole antifungal compounds, leukotriene D4 antagonists, herbicides, pesticides, phthalate esters, phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol and their corresponding metabolites, acetaminophen and its corresponding metabolites, benzo(a)pyrene, 3-methylcholanthrene, benz(a)anthracene, 7,12-dimethylbenz(a)anthracene, and their corresponding metabolites.

3. The method of claim 1 wherein the sample is a tissue selected from the group consisting of liver, kidney, brain, spleen, pancreas, and lung.

4. The method of claim 1 wherein the test compound which elicits the metabolic response is a compound with previously unknown metabolic response.

5. The method of claim 1 wherein the test compound or molecule which elicits the metabolic response induces at least a 2-fold change in the amount of the hybridization complexes formed with at least one of the nucleic acid molecules of the sample.

6. An isolated and purified nucleic acid molecule selected from SEQ ID NOs:1-117, or a fragment thereof.

7. A method of using the nucleic acid molecule of claim 6 to screen a library of molecules or

compounds to identify at least one molecule or compound which specifically binds the nucleic acid molecule, the method comprising:

a) combining the nucleic acid molecule of claim 6 with a library of molecules or compounds under conditions to allow specific binding; and

b) detecting specific binding, thereby identifying a molecule or compound which specifically binds the nucleic acid molecule.

8. The method of claim 7 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, artificial chromosome constructions, peptides, and proteins.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
CUNNINGHAM, Mary Jane
ZWEIGER, Gary B.
PANZER, Scott R.
SEILHAMER, Jeffrey J.

<120> TOXICOLOGICAL RESPONSE MARKERS

<130> PA-0010 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/141,825; 09/172,711; 09/172,108
<151> 1998-08-28; 1998-10-13; 1998-10-13

<160> 117

<170> PERL Program

<210> 1
<211> 259
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700416862F7

<400> 1
gtggcggcga tttctgcgtc gagcatttgg agtttcttcg ctgctgaacg ggtagactaa 60
acggcggctg acatggtgga ggaggtacag aagcattctg tgcacacact agtgttcagg 120
tcattgaaga ggacccatga catgtttgtg gctgataatg gaaaacctgt gcctttggat 180
gaagagagtc acaagcggaa aatggcaatc aagcttcgta atgagtatgg ccctgtgctg 240
catatgccta cttcaaaag 259

<210> 2
<211> 295
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 212, 227, 229, 232, 240, 243, 245, 250, 257, 267, 269, 273, 277
<222> 288, 290
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700502957F6

<400> 2
gcctcttcca ccatccggcc tagtcactgc aggggccatg cctacctatt ccactcaact 60

tgttacctct gcggctccag gcagggctta gtccaacctg cccagacacg gttcaccttt 120
ttatgcccaa gctttcgggg tgctgaggta ggggctgcct tcctgcaccc ccaaggagca 180
gacactcaag aatggagtca gctaggaacc cngggagctg cctcatnang cncttgatan 240
cangnacacn tttgcanctg cagaccntnt tcngaanaac nttgccangn tcaac 295

<210> 3
<211> 273
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 56, 220, 235, 237, 239, 249, 256
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700504502F6

<400> 3
tgaccctgct tgctgcaggt gactggtcaa gtgcgagcta gcttagttag tgtggngtat 60
aaggcgccat cattcctcca gtaagcctcc atcccaaagc aactgaggct gtggcagtga 120
tgccagcaac ctgtgtcacc caaaattatc cagccctcca cgggcactgc ctaggacctg 180
gggagggaag ggactttgca tcacatagcc tcaggttcgn gtttggctct ggtangngnt 240
gcctgaaant ggtggnttcc agctggtgta cgg 273

<210> 4
<211> 264
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700505170F6

<400> 4
ggctggtctg cgatggcccg ctacctcggc tcgctggaac catgtgtggg tccggcactt 60
gagactggaa tcctgaaagg ggtgaacctt cagcggaaac ttgcggcaaa ttttactccg 120
tccggacagc cacggcggga ggaggcagtg aatgctttgt gctggggcac aggcggcgag 180
acccagattt tggtgggatg tgcggacagg accgtgaggc actttaatgc ggaggagggt 240
acattccaga ccagagatac tgcc 264

<210> 5
<211> 268
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 231
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature

<223> Incyte template ID No: 700511170F6

<400> 5
ggcaacaaga cgctgtgatt ggaagcaatg acgaagtcct cacactccgg ggagtgggta 60
tgtgctgcta catcatgtga tgggcagcct ggaggggatg cagggcgcct ggagctatgt 120
ccagggtggc atgggtgccc tctcagatgc cattgcaagc tcggctactg cacatggagc 180
aagtatcttc acagagaaga ctgtggctaa ggtgcaagtg aacagcgaag ncgtgtccaa 240
ggggtcgtgc tcagggccgg cgaggagt 268

<210> 6
<211> 284
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 28, 37, 44, 67, 71, 88, 102, 172, 238
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700607827F7

<400> 6
gtgtacctac ctgctgagga aggaaggngt ggatggnacc cagnaacctg atgtccagca 60
caagggnaac nggcgtggct ttctaatnta caagcctggg tnactacagc ttgcagctac 120
ctaacccatg caggaggcga accctctgag cccagttgct attgtgacca tnaagatgtc 180
ttgccacaca gcttccaccc agtctgggtt taatgggaag ttacctaacg attacccnca 240
gaagacacat gagacgcttg cttcaaagct ctcggatgca gcca 284

<210> 7
<211> 243
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700500814F6

<400> 7
actttttcta atgtcttatg gccacttctt atgagaatgg ggagctgctc tgcctgaggg 60
tcgtgagagg aagcgccaga gcaggcccat catcccaacc cttggccttg gcccttcccc 120
ctagctctgc agcatttctt cagatcctct ttcctgagag tcaaggagac taaacaccaa 180
taaaccagac acaaccttcg tggccccaaa ggagaaaccg attagagggt tctctgctag 240
atg 243

<210> 8
<211> 259
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 235

<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700526819F6

<400> 8
ggcgcaggcg gaaggggcct gtcaccgtcc gctgcgacgt cgcggctgga gttgaacctg 60
gtgccggctg ctttgcgctg tgagtcgatg gcggccgaaa aacgagaacc ggacgagtgg 120
cgtctggaga agtatgtggg tccctggaag acatgctgca gccctgaaag tccaagcaag 180
taaaccgcct cggaagtgat cagtgagtac tcccgcaaag tgactttctg aaggnatgct 240
gaggctgaga agtgactct 259

<210> 9
<211> 255
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 2, 4, 53, 55, 86, 88, 92, 106, 133, 211, 214
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700528176F6

<400> 9
cnanggccct cagggaatca agaggagcca gcctgatccc tggcctctgg agncntaaaa 60
caagtgtgtt tttgcaggta gcctangntg gntgtcgatg gagctncagc ctgcatggca 120
ttaggcagga agncactctg gatgattgtg cacatgagaa cctagtcagg gagggagggt 180
ttaaggagag gcttagaata caagtgagaa ncancgagaa agggaccaag tcctcagaat 240
agaagctatc tgcct 255

<210> 10
<211> 269
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 30-31, 140
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700528082F6

<400> 10
ataaaagtga aaactgggca agggcaaggn ngctgggcgt gaaccgctta ctagataatg 60
gtctctaaaa attggctctg aaaaccctgt ttgtgtattc gttttatgag tgcttaaaaa 120
tggtgtgacc agggcatggn cactgtcatt ggaacagcaa catgcttgct ggcacattgg 180
aatggggaaa tgtgaagaaa gctggcatca ggcctgcggc acccatttct ttgatgaaag 240
tgttgtgtca aacccccact aatcatttt 269

<210> 11
<211> 254
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 36, 47, 67, 80, 82-83, 92, 111
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700535328F6

<400> 11
caaggcgttc tgctgcgaga acgacattga catcgngcgc gtggtanacg tgcggaggct 60
ggcggcnatc gtgggcgccn annacgagag gngcgcgccg tgagacttgc nttgcatcct 120
catttcgaac cctaatgaag acacatggaa ggaccctgcc ttggagaagc tcagtttgtt 180
ctgcgaggag agccgcagct tcaacgactg ggtcccagca tcacccttcc gagtgacagc 240
ctgcagggac cttg 254

<210> 12
<211> 244
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 50, 102, 169, 171, 204, 210, 234, 242
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700368973F6

<400> 12
cacaatccca aactggaaaa acttaaaaag gaatcctgct gtgaaaggtn tatattactc 60
tagatttttc ttactgtaaa tattgtaaga ttgtaatact gncaatattt tattaaccaa 120
caaatgttaa tctatgtgaa atcagactta tttaaagggc tgctattang ngtgtggccc 180
tttgctgaca gattaagtat attntgagtn agataactta ttaaggatgg aacnttaaag 240
gntc 244

<210> 13
<211> 237
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 11, 20, 132, 136, 149, 151, 158, 171, 173, 184, 187, 190, 193
<222> 204-205, 208, 212-213, 215, 221, 227, 230-231, 236
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature

<223> Incyte template ID No: 700368974F6

<400> 13
gtccttagct ngtgggcggn ggggtgcagt ctgctactat ctgaacgaat ttaatgtggg 60
agcatgcctt atacaacaca ggaaattaat gtgtgatcta atgcgtgatc tatgacttat 120
tacaatacag anttangtgt gaacctgcnt ncaaaacngg tcagaatttt ngnaatggcc 180
ggantgnacn ggntgnttat taannntgnaa gnngntggga naggccnggn ntgcgnt 237

<210> 14
<211> 235
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 38, 49, 53, 55, 107, 125, 132, 136, 138, 150-151, 163, 165-166
<222> 179-180, 184, 186, 189, 192, 199, 219-221, 223
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700369461F6

<400> 14
ctttaaccgg tgggctgctg taagaatcgg tggcaggnct ctctctgcng ggngntaatt 60
gctctggaac gctactagga cccgaatact aaggccacat ctctacngtc taagagggga 120
aatangatag cnttgntncc acatgtggcn nagtggggtt gcngnntatn gcttaacann 180
tacnanttnc antgattant gtggtggtaa gatggcttnn ntnaaaactg ccgcc 235

<210> 15
<211> 205
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 17, 19, 21, 56, 69, 75, 81, 103, 109, 111, 122, 127, 130, 137
<222> 148, 151, 153, 166, 174, 177, 190, 193, 195
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700782579F6

<400> 15
gggtcattta caacctntna nacaagggga cgcccccaga catgccagtg ttcacngaac 60
aagatgcang tccancagga ncagatagac tcagtccatg gantggctna ncaaccagcc 120
cngggcngcn caactgntgg acangtangg nantttctgc gacaantggt gagnttnttg 180
cgaagctccn tgncngaaag aatgg 205

<210> 16
<211> 236
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 11-12, 33-34, 46, 76, 85-86, 103, 123, 127, 144, 152, 157, 162
<222> 169, 174, 183-184, 186, 191-194, 217, 224, 230
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700782665F6

<400> 16
ggtctccatg nngaccgggc gccttggggt tgnngagacg ctgcangccc ttaacgcccg 60
cttgtagggc ccttgnaacc ccggnntggt ttaaggaaaa cgnatgcccg acaccttcgt 120
gtncganact ttttggcacc gcgnaaactc gnttacngtg gnttttnang tagngggtat 180
gtnncncgag nnnntttagg ccggcttgtg ctgcggnatt gggntgaaan tgtctg 236

<210> 17
<211> 267
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700480927F6

<400> 17
ctggaccaac atcacaagaa atgaataaag cagatttctc tgttgagttc tgcagtaaac 60
cacctaaaag ccaatgtcaa gtcagccgca gacttactta gcctgcctag cactgtagag 120
ggacttcaga agagtgtcgc ttccattggc aatacgttga acagtgtcag ccttgctgta 180
gaggcaatac agaagaccgt ggatgaacac aaggcacctt ggagttactg cagggcagtg 240
tggagaccaa tggaagcaac caaatca 267

<210> 18
<211> 271
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 2, 92, 162, 181, 247-248, 253, 256, 262, 269
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700481732F6

<400> 18
gncttattta tgtatgaaaa tgcagaaatc tgtacattcc tcaagccagt cctgtcgagc 60
caggtctgtc ccatccttgt acctcaaccc antcccacct ggcctgaaca ccccatgaga 120
cagagctggt ctctgggctg gggcccccag gcctgggctg gncaggcaga ccctaccccg 180
nagtccactg gctccagtct ccgaggctct cctgggctac aaagggggac cacacacacc 240
cagaatnntt tantgnattg gngggcccng g 271

<210> 19

<211> 283
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 37, 131, 199, 205-206, 208, 232-233, 239, 244-246, 251, 253-254
<222> 261-262, 266, 269, 274-275, 280
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700483143F6

<400> 19
ctaaaattaa gatagaagtg aatgagacag atatctngta agacactgta ttttcttgtg 60
tgatcagatc tagtgtggtg ggatgataga agttgaactt gctttattgc tatgggttaa 120
aatattttgt ntcattaaaa tggcctattg aaatgctttt ctgttcctat aataaaataa 180
cctgatgaaa aagtaaaana aaaannanaa aaaaaaaggg gcggcccccg cnnagggcnt 240
tttnnncccg ngnnttantt nngccnggnc cctnnggggn cca 283

<210> 20
<211> 256
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700484538F6

<400> 20
ctggcctcag cttcctaagt tctgggagtc ggacaggtgt ttgccacaca catggcccac 60
cggggaccta gaacctacag tgaaccgtca cccaggctct gtggatgttc tgcatcctga 120
ggtagacagc ctctaatatc ctgttaggga cctaggacca gagctggggt gcccaggcat 180
gtcccaacat gtcgcatcgg ccacagggat atcggttgaa gtgcatttgg aagtgtgctg 240
ggacgccagc cagctt 256

<210> 21
<211> 272
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 64, 68, 124, 140, 145, 213, 260, 267, 269
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700480077F6

<400> 21
ctgacctgac ccatgattta aggaccgtag tttagcacgg accactgcaa aggcgggcta 60
aggnctgntg ggctaaaggt ctctttgagc ccagtggcta tagtcacacc ttctttgctc 120
tggnccagga ggcctacttn ttctntactc gtggaatcct ggaatcttaa agataaaaga 180

acctagaaag aaaatcaaac ccactttcct tgngggcaga tggtaatatg ggactgagac 240
agcaaacctg gggcttggan aggaccnanc tc 272

<210> 22
<211> 270
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 40, 44, 46, 114
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700060207F7

<400> 22
taccctgca tgggataccg tttctcgacc ccagtgcacn tggntnctgt catcccatga 60
aattgcagca agggcagtct cttttgtggg aacagattaa ctcctacaca tgangtagat 120
tcaacacctg ccaggaaagc agaagcatta cttaagtgtc ctgtgaaggc aaacatcaag 180
tcaattcagc ttatcttgaa gagtggcaaa ccatgaactc caaatgtcat tgtgtgaaac 240
tgaacgatgg tcatttcatt ccggtgctgg 270

<210> 23
<211> 250
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 45, 49-50
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700133105F6

<400> 23
ccctcctgta tctgaaccca gcttctcagc tctgagatga gtgcnggann ggcttcccaa 60
cctatgctca ataccacagg cagcctgcag gagggagaaa tgggtaaaat gttccatggg 120
aaatgtctca gaatcgtctc ccccgaatct cctgctaagc tttactgctg ctatggagtg 180
atcatggtcc tcagtgtagc tgtagttgct ctttctgttg ctttgtcagt aaaaatgaca 240
ccacagatct 250

<210> 24
<211> 226
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700138117F6

<400> 24
gaggattcac tcacatttgc ttcccgctgg ccatgagtga gctgcccttt ctgagtccag 60
agggagccag agggcctcac aacaacagag ggtctcagag ctccctggag gaaggctcag 120
ttacaggctc agaggctcgg cacagcttag gtgtcctgaa tgtgtccttc agcgtcagca 180
accgtgtcgg gccctggtgg aacatcaaat catgccagca gaagtg 226

<210> 25
<211> 265
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 85
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700268788F6

<400> 25
cggaggtgct cccaggcggc tgcactggct cggaggagta gcaggaggag ctccgcgcag 60
gaacaaacct ggaggcaaac caganggagg caatgtttga atgactgtaa gaagaccaga 120
cagtgaaaat gtcagccctc aactggaagc cctttgtgta cggagggctg cctccatcac 180
cgcagaatgt ggtacatttc caattgattt gactaagact cggcttcaga ttcaaggcca 240
gacaaatgat gccaagttcc gagag 265

<210> 26
<211> 257
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 58, 99
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700270924F6

<400> 26
ctgggatccc cagggctaat gggcatcctg ttcttgcagc agggcactgt gagaaagnct 60
ctcaccgtga ccaagtttct ctgagtgtcc agccaaccna ggctcaccag ctcctccag 120
ctaccgcccg tccatcaggt cagctgccaa ccccaggctg aacaccaacc ccagctatga 180
gctcctggag gcatgactcc ctcagggcca gcagctccga tcccctccca gtagttatca 240
ttggcaatgg ccctcgg 257

<210> 27
<211> 244
<212> DNA
<213> Rattus norvegicus

<220>

<221> unsure
<222> 2, 8, 50, 56, 63, 76, 177, 219, 233, 240
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700303722F6

<400> 27
gncgaggnca ccaaggtgtt tctgcctcta cttagaaagt ccaaggggan gctggntaac 60
gtnagcagca tgggganccat gattccattt cagatgatgg ccgcctacgc ctgcacgaag 120
gcagctataa gcatgttctc agcccgtcat caggcaagag cttccaaatg gggagtnaaa 180
gcgggaccat cattctggag cttcaaacca acatcgtang ctcacaggac agntgggatn 240
aaag 244

<210> 28
<211> 263
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700502047F6

<400> 28
ggaaatgact ggtctgaagc ggcttggcag cctgagcagt caggtactgc ggctctactg 60
gcactgcctg agggtccaag gactgtggtg attctcatgg aggaccctga gatttctgca 120
atctgatcag tgtcaaatgc cactggattc gctctgagac tcttgcccta gaggatggcc 180
aaagggctcc tgatgaccta tgccctttgg cttttggggc cctgttggac tacaccacct 240
gtatctggga agggacagcc atc 263

<210> 29
<211> 259
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 52, 137, 167, 189, 235, 254, 256
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700502260F6

<400> 29
ggactacccc caccatgccc gtgtcattga gactctgatt gtccactatg gnctggtctt 60
tgaggaggag ccagaagaag cagctggcag ccaagagggg gcgtccgcca gtgtgcccag 120
ctggagactg ctgaggncat tgtcttcccc cagcaggagg aggcggncga tggaaaccga 180
gaatcccang tgcatcaatg actcagactc agagctggaa gaggcttctg acctntttcg 240
cctcggacgc cacncnctc 259

<210> 30
<211> 260

<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 140, 158, 161, 170-171, 177, 191, 193, 199, 203, 206-207, 212
<222> 217-218, 220, 224, 226, 231, 238, 243, 246, 248, 253, 255, 258
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700142213F6

<400> 30
cacagggttt tccccaagag cacgcctctt tacttcaggg aattctcgga aactttcagg 60
attcgggggg cttccctggg aggagacgag agaggattag aagcggacat ccacggcttc 120
ttgtgatgac cacgcctttn gtctttgcta gaactctntg ngacttcccn nggtganttc 180
taatcacgga ntncacgcna agnttnngga antacgnngn cccntnaaga naaacacntt 240
ttnggngngg ggnanaanct 260

<210> 31
<211> 288
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700229555H1

<400> 31
caggagtacc actcacaggc cacctggcag gaagagataa gcccccagcc cccgacatcc 60
aggacgcccc gaacctgcca atgtgtgtag ctatacctta ttacctcatc atgtgaaata 120
gccaatcata tgtgaacatg tctatgtgcc tcgtttgaat ccaccaatcc ctgtaactat 180
gcatctgctt ctgtacgcct gcttctgctt ccccaatccc tataaaagcc ccatgctgga 240
gctgctgggc gcgcaagtcc tcctaagaga ctgtgtgccc gcagtacc 288

<210> 32
<211> 258
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 53, 244
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700626839H1

<400> 32
gggtgcgcgt ggagttgcgc atgcgccttc ccgccgcgca gggcaaaggt ggnggcgctc 60
tggtgaatgg ttggttgctg tgcaagagcg ttttctggct tttggtggcg aaggcggcct 120
ggccgcgagg tgcagctgct ggtgggcagg tgtactaatg tctacagact atgagctttc 180
agaatctctg gagagagtac aaagttctgc atgttatggt acctttaatc gggttcatac 240

attngggtgg cacagaat 258

<210> 33
<211> 268
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 7, 11, 40, 53, 150
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700627890H1

<400> 33
gtacaangag ngccggggct tgggtctagt tggaggggan gcagtggcca gtncagggct 60
cagatgagag agttagccga gttaggggca gctactagga tgggggcagg aggagaagcg 120
gggctaacta taaagaagac tagatttcgn cacagtgggt atgtggaagg cagctttcaa 180
accgcccttg tcaaacaaca cagggccagc agccttcaag accaggctat ccctgccgtc 240
tgctggcatg ggggcacttg taccgtcc 268

<210> 34
<211> 299
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700059105H1

<400> 34
tcatcacaac ccaactgtgt ggccaaatcc agaggtgttt gacccttatc gatttgcacc 60
agagtcttcc cgacacagcc actcattcct gcccttctca ggaggagcaa ggaactgcat 120
tgggaaacag tttgccatga atgaactgaa ggtggccgtg gccctgaccc tgctccgctt 180
tgagctgctg ccagatccca ccaggatccc aatccccata ccaagactcg tgttgaagtc 240
caagaatggg atctacctgc gtctcaaaaa gctccaataa tcttgacagc acaagacag 299

<210> 35
<211> 300
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700059610H1

<400> 35
aaacaacctg actttcttgc gtgtgaggag tgccttttat gggaacagca tcatctacaa 60
tatgtcctct gatggccgtt tgtcccgccg ggcctgccag attgctcatg agcacacaga 120
tggagtgatc aaaatgagga aggctcagct gcagaatgag gaagagcttc agaaggccag 180
gaagaagagg cacttggatt tcctggacat cctgttgttt gccaaaatgg aggatgggaa 240
gagcttgtct gatgaggacc tgcgtgcaga ggtggacaca ttcatgtttg agggtcatga 300

<210> 36
<211> 296
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700059630H1

<400> 36
gggtttctct gtatttaccc ctacaagatc cctggatggt gtctctgggt tcttccaagg 60
ggccttcctg ctcagtctat ttctggtgct gttcaaggca gtccaattct acttacgaag 120
gcaatggctg ctcaaggccc tcgagaagtt cccatccacg ccttcccact ggctttgggg 180
ccacgacctg aaggacagag aattccagca ggttcttacg tgggtagaga aattcccagg 240
tgcctgctta cagtggctct cagggagcaa aacacgagtc ctgctctatg accctg 296

<210> 37
<211> 286
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 204
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700062959H1

<400> 37
ggcccatgga gcacacccag gctgtggact atgttaagaa gctgatgacc aagggccgct 60
actcactaga tgtgtggagt aggagctacc accctcccac ccctcgctcc ctgtaatcac 120
ctaacttctg ccgacctcca cctctggtgg ttcctgcctg gcctggacac agggaggccc 180
agggactgac tcctggcctg agtngtgccc tcctgggccc ctaagcagag tccggtccat 240
tgtatcaggc agcccagccc caaggcacat ggcaagaggg attgac 286

<210> 38
<211> 289
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700606459H1

<400> 38
ggtgagtccc gtgtggagaa aatatacaag taagaccgct acgtgcctgg cgactggaga 60
tgtgatgggg cacagcgcac agagagccat aatggcctca tcgtacaggt ctgggacgct 120
cagcaacacc ccagcaggca ctgcactgtc tagtggacaa gctctgttag caggaagagc 180
ttctctgcgt ctgtccaaga aaggctggtc aaggctccct accacctaca catactgtgg 240
ttggagaaat caaggttcct ggcaaaagag agtagcttca cgggaggca 289

<210> 39
<211> 79
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 39, 48, 66, 68, 72
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700608882H1

<400> 39
cttaacgctc ctgccacgcc gcctccgccc gtgcaatgnc tctgtagncg gcgatctacg 60
tacgtntncc cngccccgt 79

<210> 40
<211> 248
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 8, 21, 24, 35, 40, 52, 78, 90, 104, 137, 169, 209, 220
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700483988H1

<400> 40
tttagttnca atcatggagg nctntcaggt gaggncctan ctcccaaagc cngcgctgag 60
tctacaactt ctcctaanag gtgttccggn cggcgttggg gtcnctgcgg aggcggctaa 120
atcggccgca gtttctncca tggttgcgcc cgctgtgttg cgcgctctnc gtaagaacaa 180
gacccttcgc tatggagttc ccatgttgnt gctggttgtn agtggttctt ttggtcttcg 240
cgaatttt 248

<210> 41
<211> 352
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 28, 324, 337
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700501033H1

<400> 41
ctctctgcct atgttctgag gttggagnct ttattacaga agctggtgca gaaaggagca 60

attgagaaag aagttgtgaa tcaggcccga ctagaccaag tcattgctgg ggcaatccac 120
aagtcagttc gaagagagct tggactgcca gaaggtagcc ctgccccagg cttattgcag 180
ttgctgacac tgataaaaga taaggaggca gcagaggaag aggtccttct tcaggccgaa 240
ttagaaggac atttcacttg acccaagacc agcaaggctg tcatgagcag aacatgatgg 300
aggagctcat agaagtgatc agcncatccc ctttggnctg ccaagtaatt gc 352

<210> 42
<211> 233
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700137747H1

<400> 42
gtttctccat agcctcagac cccacatcag tatcctcttg ctacttggag gagcacgtga 60
gcaaagaggc taaccatcta atcagcaagt tccagaagct gatggcagag gttggccact 120
tcgaaccagt caaccaggtg gtggaatcgg tggctaatgt catcggagcc atgtgttttg 180
ggaagaactt ccccaggaag agcgaggaga tgctcaacct cgtgaagagc agc 233

<210> 43
<211> 243
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 33
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700141708H1

<400> 43
tgggcagaaa ggaagccctg cagagcatca gangcccagc tagagggaca acacagagga 60
gtaatttgct gacagacctg cagggatgga cctgctttca gctctcacac tggaaacctg 120
ggtcctcctg gcagtcgtcc tggtgctcct ctacggattt gggacccgca cacatggact 180
tttcaagaaa caggggattc ctgggcccaa acctctgcct ttttttggca ctgtgctgaa 240
tta 243

<210> 44
<211> 295
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 286
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature

<223> Incyte template ID No: 700302454H1

<400> 44
gcgggccgtg ggtgatctgg tcggtaccgg agagcgcagg ttgtatcacc aacatggggg 60
actctcacga agacaccagt gccaccatgc ctgaggccgt ggctgaagaa gtgtctctat 120
tcagcacgac ggacatggtt ctgttttctc tcatcgtggg ggtcctgacc tactggttca 180
tctttagaaa gaagaaagaa gagataccgg agttcagcaa gatccaaaca acggccccac 240
ccgtcaaaga gagcagcttc gtggaaaaga tgaagaaaac gggaangaac ttatc 295

<210> 45
<211> 286
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700304405H1

<400> 45
cggaagtgaa ccaaggcact gagcggcatc taatgcacct ggagttggac atctcagact 60
ccaagatcag gtatgaatct ggagatcacg tggctgtgta cccagccaat gactcagccc 120
tggtcaacca gattggggag atcctgggag ctgacctgga tgtcatcatg tctctaaaca 180
atctcgatga ggagtcaaac aagaagcatc cgttcccctg ccccaccacc taccgcacgg 240
ccctcaccta ctacctggac atcactaacc cgccacgcac caatgt 286

<210> 46
<211> 311
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 299
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700306096H1

<400> 46
gataggaaaa taattttatt taggtttttt aaaaaagtta actttcacat ataaatttag 60
acttaaagat tacagtgtat attttccaaa aggagcgccc ctgaagggtg gccagacaag 120
ctcgccgagt gggcacaggg acactcgctc cagaaggagc tcaggtggaa gcgctttctt 180
taatcttcca cagtggccct tccctgttcc tcaccgggcc tatgactggt aagaaaaccc 240
acaaccatca tttggggcaa cagcatctca ctagatggga ataagaacat gtctaggang 300
aaagcacaag c 311

<210> 47
<211> 307
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature

<223> Incyte template ID No: 700325693H1

<400> 47
gtgccctcac gcagcttaat gtggcctttt cccgggagca ggcccacaag gtctatgtcc 60
agcaccttct gaagagagac agggaacacc tgtggaagct gatccacgag ggcggtgccc 120
acatctatgt gtgcggggat gctcgaaata tggccaaaga tgtgcaaaac acattctatg 180
acattgtggc tgagttcggg cccatggagc acacccaggc tgtggactat gttaagaagc 240
tgatgaccaa gggccgctac tcactagatg tgtggagcta ggagcttacc aacctcccac 300
ccctcgg 307

<210> 48
<211> 300
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 21, 49, 199, 226
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700059571H1

<400> 48
ctgtgctcct gagtgcaagc naggcattcc tccaagacac tgcgggtcng agcagggact 60
gttcacgctg gtgccctgtg aactctggtg gaggtcagcc aacagctgct gtgtctgagt 120
tgctgagagg agagagaatg gcttgcactg agttttcttt ccacgtgcca agtctggagg 180
agctcgcaga agttttgcng aaggggctaa aggacaactt tgctcntgtc caggtctctg 240
tggtcgactg cccagattta acaaaggagc catttacttt cccgtaaaag gcatctgtgg 300

<210> 49
<211> 314
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 13
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700060610H1

<400> 49
gcaagattga ggnggagaag gacaacctga agtctgagtt ccatctggag aacttggctg 60
tctgtgggtc taacttgttt acggcaggca ccgagacaac cagcaccacc ctgagattcg 120
ggctcctgct ccttatgaag tatccagagg tgcaagccaa agttcatgag gaacttgacc 180
gtgtgattgg acgccaccaa ccccccagca tgaaggacaa gatgaagctg ccttataccg 240
atgctgtatt gcatgagatt caaagataca tcactctcct tccttccagt ctgccccatg 300
ctgtggtcca ggac 314

<210> 50

<211> 312
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 78
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700063128H1

<400> 50
cggtcggtac cggagagcgc aggttgtatc accaacatgg gggactctca cgaagacacc 60
agtgccacca tgcctgangc cgtggctgaa gaagtgtctc tattcagcac gacggacatg 120
gttctgtttt ctctcatcgt gggggtcctg acctactggt tcatctttag aaagaagaaa 180
gaagagatac cggagttcag caagatccaa acaacggccc cacccgtcaa agagagcagc 240
ttcgtggaaa agatgaagaa aacgggaagg aacattatcg tattctatgg ctcccagacg 300
ggaaccgctg ag 312

<210> 51
<211> 248
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 64-65
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700139656H1

<400> 51
caaaatacaa caaggagata aaagtcatac agtttgtgct gctggcttat tagctctgca 60
tggnngaggg gccatggtaa gttgccaagg cattcagata taaatgtaat taagatgcca 120
tgtttgcttg cagtaatgaa gttataatca gaaactgcta aagtatgata aaaacagtga 180
ttgtttatgc acttatggaa gacaaagtga agtgatgtgg tttcttcaga acaggtgatg 240
cactgagg 248

<210> 52
<211> 115
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700141348H1

<400> 52
gtaaaagatg tggtccagag aacgtaggaa catgcctgga gagacctaat gtgctcttgt 60
tctgcaaacc catgggcatt atttccctct ccgctcaaga gctcatactg gaagc 115

<210> 53
<211> 249
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700182318H1

<400> 53
agagtttcct tttgctcccc aacctgtagt tctaagttca acaaaacagt catcaacaaa 60
agtgacggag gcttccatca gtgtcagggc tctgtgcaga cccagaatcc ctgttcccta 120
tgtcatgttc cagcattgta tagcacggtt ccatgtcaca aacagaaagg tcaggaacac 180
tgaggtctgt gaatgtcact gctgcagcga ggtcatgtca ctcctctgtc tactctgtca 240
gtgtcttac 249

<210> 54
<211> 296
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 22
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700225376H1

<400> 54
agcaagccta tttctgactg gnctgctgtg cagaatctag accactggca gtgggtgaca 60
gcccagttga ggttaatcga agtctcgtcg caggctctgc tgtaagtctg gcctcttggc 120
ctcacatctt ctttgtggga tccttcccta tctccagctt cctcagctgg tcagggagat 180
ttggtccaga actagaagcc ttaataatct gagcaggtaa gagaggagta aaatgtacag 240
tcttggacat tgactaaagg gtcctgcaga ggatatcaag gtaagtggct tggagg 296

<210> 55
<211> 169
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700225757H1

<400> 55
gctttctggg caagtctatg ttgccctagc tgacccagaa tttgctatat agactattct 60
gtctcaaact cacagaaatt ctcctgcctg tgcctcctga gcgagcacca ggattaaagg 120
cgtgaatcgc tgtccccgtc ttttttcttt cttcttttaa taacccact 169

<210> 56
<211> 191
<212> DNA

<213> Rattus norvegicus

<220>
<221> unsure
<222> 190
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700268788H1

<400> 56
cggaggtgct cccaggcggc tgcactggct cggaggagta gcaggaggag ctccgccgca 60
ggaacaaacc tggaggcaaa ccagaaggag gcaatgtttg aatgactgta agaagaccag 120
acagtgaaaa tgtcagccct caactggaag ccctttgtgt acggagggct ggcctccatc 180
aaccgcggan t 191

<210> 57
<211> 249
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 45, 118, 128, 163, 245
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700270924H1

<400> 57
tgggatcccc aggggctaat gggcatcctg ttcttgcagc agggnactgt gagaaagtct 60
ctcaccgtga ccaagtttct ctgagtgtcc agccaaccca ggctcaccag ctccctcnag 120
ctaccgcncg tccatcaggt caactgccaa ccccaggctg aanaccaaac ccagctatga 180
gctcctggag gcatgactcc ctcagggcca gcagctccga tccctcccag tagtgatcat 240
gggcnaggg 249

<210> 58
<211> 294
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 19, 32, 35, 104, 131, 188, 220
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700291661H1

<400> 58
actagcttct ccagctccnt ctttcccgag angcngcagg gacctcggcc tccagcttac 60
cgggcggatc gaagcagcgg tcgggatggt actgctgggc ttgntgcagt caggcggctc 120

ggtgctcggg naggcgatgg agcaggtgac aggaggcaac ctgctttcca cgctgctcat 180
cgcctgcncc ttcacgctta gccttgtcta cctgttccgn ctcgcagtgg gccacatggt 240
ccagctgccc gctggagcga aaagtccgcc atatatttac tctccaattc cgtc 294

<210> 59
<211> 304
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700301979H1

<400> 59
gatatattaa tcaaaaagaa aggaccacga ctcatgacct cccatcttcc catgcatctt 60
ttctccaagt ctctcttcag ttccaaggcc aaggtgatct atctcgtcag aaatcccaga 120
gatgttcttg tttctggtta ttatttctgg ggtaattcaa ctcttgcgaa gaagccagac 180
tcactgggaa cttatgttga atggttcctc aaaggaaatg ttctatatgg atcatggttt 240
gagcacatcc gtgcctggct gtccatgcaa gaatgggaca acttcttgtt actgtactat 300
gaag 304

<210> 60
<211> 293
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700302770H1

<400> 60
gtagccactc taactagggg cgtgctgaga caagaccacc tcattcctct gctgcttttc 60
agacaggact gtcctgccga cccaccatga tccaggctgc actgttcctt ggctgtatct 120
tactgtcctc ggtgaccgcc tttccatgga agactcagga tggtggcctg ccccatcagc 180
cagctggcac agaaactgag cctacacaac tgctctacag caagagtcct cctccgacct 240
ccagtacctg tcggaacctc ctaagcatgg cgcccctgcc ccctgtagtc ctc 293

<210> 61
<211> 174
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700303111H1

<400> 61
caacagcaaa cgtgcgacca actcttcggc tagtgaatct ctaagccgcg aagagtgctt 60
tgaagtagct ttaggtggaa gatgtcagaa agtaactcgg cagagggtag cgacagaagc 120
gaggagcagg tgtctggtgc taaagtcatc gcccaggccc taaaaacgca agat 174

<210> 62
<211> 273

<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 52, 139, 249
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700303390H1

<400> 62
ctcagatggc aggcatggca tgcatgggat cttgttccct gagacaaagg cngatgcaga 60
gggcatgtga ataaatcatg aggggcccac agcaggccag caggccatag ctgacctcat 120
tctggaagtg agagttgang agaccccagc tgggacagaa aaggtaccac gcctataacc 180
atggcctaac cgagggccag cagtggcagc ctccctgaaa gggacttcca gtccatccac 240
aggcaccgna gaaccagcaa gacatagcca gcc 273

<210> 63
<211> 279
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700303722H1

<400> 63
gtgactctga gtgtttgagc aggtaacttc tacctttgca cctctatcgc aacaggtcca 60
aaggttccaa aggagctggc aggacactca gacaagatcc actggcttca ggtgtgccta 120
gtcctggagt tcagaaagac ggaggcagct gaatgtggtg ctgaaccaac aacatctagc 180
tacaagggga gccactcctc cacccagcga ctgtgactgt tctcacaggt ctgaatttcc 240
tgttggtatt cacaaagatg ctttttattt ttaacttct 279

<210> 64
<211> 275
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700306343H1

<400> 64
gagaaaggcc accacctagc taggtgaggt gtgccagcat ggtcctgggg gtctcactgt 60
ccccagccct gggacgctgg ttccgccatg caatcccttt cgctatcttc acgctgttac 120
ttctttatat cagtgtatgg ctcttccatg agtggccctt tgagttgcca gctcaaagaa 180
ctcagcagtc cggcctgtgg gaactcaagc tctcttctcc ttctccagcc ctcacctctc 240
tgcttcctgt cacctcaggt gttttacaag gctga 275

<210> 65
<211> 294
<212> DNA

<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700306615H1

<400> 65
catccgtggg ctggctcacg ccattcgcct gttcctggag tatacagaca caagctatga 60
ggacaagaag tacagcatgg gggatgctcc cgactatgac agaagccagt ggctgagtga 120
gaagttcaaa ctgggcctgg acttccccaa tctgccctac ttaattgatg ggtcacacaa 180
gatcacccag agcaatgcca tcctgcgcta ccttggccgg aagcacaacc tttgtgggga 240
gacagaggag gagaggattc gtgtggacgt tttggagaac caggctatgg acac 294

<210> 66
<211> 283
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 2
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700330359H1

<400> 66
gnggcccaat ggcccctgag taaaaggtgg tcactgagag tcctaaggcc cgagtaggaa 60
catgcggctt agagccagtg tcgtgaccca gaacgtccac tcttgtacag gtagatgagg 120
aggtgttcgg gtgcccgcag gcggtatccc gcctggcttt cgccctagcc tttctgcaac 180
gcatggacat gaagccgctg gtggtcctgg gactgccggc cccgacggcc ccttccggct 240
gtctctcctt ctgggaagct aaggcacagc ttgctcagag ctg 283

<210> 67
<211> 263
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700368974H1

<400> 67
tccttagctg ggggcggggg gggcagtctg ctactatctg aacgaattta atgtgggagt 60
catgccttat acaacacagg aaattaatgt gtgatctaat gcgtgatcta tgacttatta 120
caatacagga atttaatgtg tgagccatgc cttcaaaaca tgtctagaat ttctggaatt 180
ggccggaagt caacagggat tgcttattta acctttcaaa tcactcattg tgactagggc 240
acatggtctt gcgcttgcta tga 263

<210> 68
<211> 269
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 129, 133, 153, 163, 165
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700373118H1

<400> 68
gtttgaagca gggcgtaggg aaaagcggga ttaaagagta ctacctttat tagctccttc 60
cctccagaat aggtgcaaat ccctccccat gcccatttcc tgccacctgg ggtaaggatg 120
tggcactgnc agnctgtcag cccactgact ttnagtcttc agntngcagt ctgggcaaat 180
accagcgagc tctgttgaac caagaccagg ccttcagagc atctgaacca ctgtggcctt 240
ctctcctcag ccttcactgt ggcttttgc 269

<210> 69
<211> 288
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 159
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700375521H1

<400> 69
gcagccatgg atcgcgggga ggaacctctg tccgcgaggc cggcgctgga gaccgagagc 60
ctgcgattcc tgcacgtcac agtgggctcc ctgctggcca gctatggctg gtacatcctc 120
ttcagctgcg tccttctcta cattgtcatc cagaagctnt ccctgcgact gagggcttta 180
aggcagaggc agctggacca agctgaggct gttctggagc ctgatgttgt tgttaagcga 240
caagaggctt tagcagctgc tcgtttgaga atgcaggaag atctgaat 288

<210> 70
<211> 280
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 16, 84, 96, 112-114, 118, 171, 226, 228, 236, 240, 267
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700461546H1

<400> 70
tgaacaaact tctganaact tagttgacag tgttttgagt caactgaaaa aagcatgact 60
tttggaatct ctgaatgcct tggntctcag tattancact ctattgaatt tnnntctnat 120
taaagtatgt agtttttag actttttcc tgacagtatt atgtaatttt ntggcgtggg 180

tagatgggag tgtcgcttgt atgttaccat acagctgaca tgtatntntt gtctantctn 240
attatcttag tagtttcatg ctgtggnatg taccataacc 280

<210> 71
<211> 271
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 25, 64, 70, 82, 102, 155
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700480077H1

<400> 71
ctgacctgac ccatgatgta agggnccgta ggggagcatc accactgcaa aggctgacta 60
aggnctgttn ggctaaaggt cnctttgaag cccagtgtct anagtcacac cttctttgct 120
ctgggcccag gaggcctact tcttcttttt ctcgnggaat cctggaatct taaagataaa 180
agaacctaga aagaaaatca aacccacttt ccttgtgggg cagatggtaa tatgggactg 240
agaacagcaa acctggggtc ttggagagga g 271

<210> 72
<211> 210
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 187
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700480949H1

<400> 72
gggcagaggt ccagggaata agggaggctt ctaccaatga ttttgtttaa tggtgcttga 60
cagagatatt gtatggttct ctgagagctc ccctgaaaac cttacctcca accacacaag 120
ggttcctccc agagagcgct cgctgggcag caaggacaca ctcccatatt gccaagcata 180
tcaagtnccc aaagattggc agaaaattcg 210

<210> 73
<211> 256
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 70
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700483259H1

<400> 73
gtgacgtaca tggaaaacaa agcctacggg gacaggctca agccgcagac agcagcaagt 60
aaagcgcctn cggccctgaa gcatggcagc tatcccttcc agcggctcgc tcgtggctac 120
ccatgactac tatcggcgta agtagcccct cgccagcccc gcccagggct ggcccagggc 180
tctgtggctg acccgcctcc ccttcccagg acgtctgggc tcctcgtcca gcaacagctc 240
cggcggaagt gcagag 256

<210> 74
<211> 259
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 219
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700483475H1

<400> 74
ctctcgtact tcggcaatgg ctggattccc accgtcatca cggcctttgt ccttgctacc 60
tcccaggccc aagctggatg gctacaacat gattatggcc acctttctgt ctataagaaa 120
tccatatgga accacattgt ccacaagttt gtcattggcc acttaaaggg tgcctccgcc 180
aactggtgga accatcgaca tttccagcac catgcgaanc caacatcttc cacaaggacc 240
ccgacataaa gagcctgca 259

<210> 75
<211> 264
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700498995H1

<400> 75
gacagtagat gcccccaaag ctctagtaga tgatagtgtg ggggctgtgt gcggctccta 60
cctgtgctgt tcattcacag tgcagtttaa gggagcaggc gccactgcat tccttggctg 120
tgccctgagg gtgcttgctg ctttatatag taacagtcaa ttaaggtttc tttcaggaag 180
agaaaaggga tggttttgag gggctcagaa aataggattc agtgtgtaac ataacaggta 240
ggttgtcggc acatgctgat atcc 264

<210> 76
<211> 271
<212> DNA
<213> Rattus norvegicus

<220>

<221> unsure
<222> 218, 228, 255, 270
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700504333H1

<400> 76
gtttattttg acacagacat ggacaaagcg atggagcgct atgtctctat gcccaaggaa 60
aaggctccag aacacattcc ccttctcttc attgccttcc catcaagcaa ggatccaacc 120
tgggaggacc gattcccaga ccgatccaca atgactgtgc tggtacccac ggcctttgaa 180
tggttcgagg agtggcagga ggagcctaag ggcaagcnaa gtgttgcntt ggaaccctca 240
aaaaaacttc ccggnaaccc tttatggggn a 271

<210> 77
<211> 167
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 11-12, 17, 21, 24, 48, 66, 72, 96, 128, 135, 162, 166
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700505040H1

<400> 77
caacaatctt nngtggnctg nctntctgga actgggcatc atcagctnat gctgccatac 60
gcctgntgga gngccgtggg gtgaaggtcg cccgtnccct ggtgggtacc ttcatgtcag 120
cactaganat gcgtngtgtt tcccttactt tgatgcttgt gnatgna 167

<210> 78
<211> 267
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 5, 18, 26, 39, 90, 92, 122-123, 132, 137, 145, 152, 160, 168, 171
<222> 173-174, 186-187, 213
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700505423H1

<400> 78
ggggncttct gtgaggcnct gatacncatc gaggctgtna ttcagccagg ccacatgaag 60
ccccaagatg ggtggctttt cctgtatgan tnagtacaga tatatccatg gccggggaat 120
tnngactggc anggtcncca gggancacca gncagctttn tcaagaantc ntnnggttcc 180
cttggnntca caggaaccta ttacctttca tgnggtctgg ggttctggat ttagggtctt 240
tgggacagtc ccagttagaa gccttgg 267

<210> 79
<211> 267
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 14, 20, 22, 24, 81, 248, 253
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700510534H1

<400> 79
tgaaagcgta aggncttgcn tntnagaggc tctgtagtga gttctgtttg cctataaggg 60
aagtggaaca accgagacac ncgcacttct ttcgagtgtt aaggagcctg ggaggagcag 120
gcagccgctt gctttgagca tgctcaggtg gggctgtccg ccgctgtggg gaaggcaccc 180
tgcagcaggg cttcctgccc cacctctcca ttgtagtagt gtccagatct cagaaacgca 240
gcttgaancc agnttcaaag gtaccag 267

<210> 80
<211> 291
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700513027H1

<400> 80
gggaggtttt aaaggccata ttgccaacct caccgaaagg tttcaggaac ccgaggaagt 60
gttaatgtac aactcaccac ttcacgccac ccgaggctga agttgacgtt gccttttaag 120
ccttttaca tacactggcc atttcagaaa attctcaaca ataatgtctg ccttcgagtt 180
taagtcatgg tgtttttag aattgacttg aaatgaaaat atcacaaagt gaatatatca 240
gctggtgatc gagtgactga aacccccctg gtctgcggtt gaccagttca g 291

<210> 81
<211> 273
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700528082H1

<400> 81
ataaaagtga aaactgggca agggcagggg gctgggcgtg aaccgcttac tagataatgt 60
tctctaaaaa ttggctctga aaaccctgtt tgtgtattcg ttttatgagt gcttaaaaat 120
ggtgtgacca gggcatggtc actgtcattg gaacagcaac atgcttgctg gtcacattgg 180
aatggggaaa tgtgaagaaa gctggacatc aggcctgcgg cacccatttc tttgtatgaa 240
agtgttgtgt acaaaccccc cactaatcat ttt 273

<210> 82
<211> 268
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 68, 121, 165, 174, 182
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700528176H1

<400> 82
caaggccctc agggaatcag aggagccagc ctgatccctg gtctctggag tcttaaaaca 60
agtgtgtntt tgcaggtagt cctagttggg tgtcgggggg aggctgccag gcctggcatg 120
ngacattagg caggaagcca ctctggatga ttgtgcacat gagancctag tcanggaggg 180
anggttttaa ggagaggact tagaatacaa gtgagaagcc agccgaggaa agggaaccaa 240
gtcctcagaa tagaaggcta tactggct 268

<210> 83
<211> 289
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700534427H1

<400> 83
aaaacaggca agactttgga gaaagcagac caggtatgat ggccactttg ccaccaacag 60
cccacacttc ccagcaacct gtaaacatag aggacgaaga tgggatcctg gatgagtatg 120
accagtacag cctggcccaa tcttatgtcg tcggtggagg tcggaaagga cgtaccaaga 180
gagaagctgc tgccaacacc aaccgcccca gccctggtgg gcatgagagg aagctgctga 240
ccaagttcca gaactctgaa aggaaaaagg cctggcgctg agacagagc 289

<210> 84
<211> 290
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 58, 157
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700535328H1

<400> 84
ggtctgacca agtgagaaag acagcagggg cacccaggcc tcagacactc tggcgtantc 60
ccaaagaaag atggccacag cccagctccc tggtaccaag ctgtcatccc taaactctgc 120
tctgtgcccc ttgtgggcag acgttaatca agccctngcc ctttctgatg ggcccctcca 180

tcccgggaac actaaaaggt agtcttactg tccaccaccc tacacctgtt ttcataagtt 240
atgcacaaat gcgaacagct gagacagaga tggagaagtt cttcgttttg 290

<210> 85
<211> 275
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 17
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700607183H1

<400> 85
cacagcccct accagcncac cctccataac tgcaccaaga ggatctatcc aacacctccc 60
tgagcaggag gagcctgaag actccaaggg aaagagtcct gaggaaccct ttcctgtgca 120
gctggatcta accacaaacc cacagggtga cacactggat gtctccttcc tctacctgga 180
gcctgaggaa aagaaactgg tggtcctgcc tttccctggg aaggaacagc gctcccctga 240
gtgcccgggg cccgaaaagc aaagaacccc ctgat 275

<210> 86
<211> 285
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700607235H1

<400> 86
ctgaagaccc accatgtctc tgctgactac tgtactactt ctctggggtt tcattctggg 60
cccagcaact gacacagcct gtatattcaa ggaagcctcg gaaaacagtc ccttgcccag 120
gccctggctt tctgccaatc cagtgccctg gatcacacct ggcctgagga cattcctgct 180
gtgccagggg acagtgcggg atgtagtctt catgctgagg cgggaaggag atgatggttt 240
cctggcgata gtccaacaga tgtttttctg gagggagctg gaccc 285

<210> 87
<211> 260
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 246-247
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700607396H1

<400> 87
ggccaccaag atggcggcgc ccagcggacg gtgcgcgagg ttcgtcagct gacttgttct 60
cggagctgtg gccgcgaccc gcttctacct gtcccgagtg accagagctc agtgaccagt 120
ccttatagtc gaaagcaggg tttttactgc tgaggacctg gacccgctgg gaggcttgcc 180
atggtaacag aacaggaggt agaggccata gggaaaaccc tagtggactc cacgcagccc 240
ctgcannccc gcttccgtgc 260

<210> 88
<211> 181
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 27, 180
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700607505H1

<400> 88
caaagaaaga aaacactcct ccgaggnccg cagcaaaggc agagaaagat ctgcaggatg 60
accttcatta cagcaggtgt tatattttat ctttttgcc tccgtttcta gtgaatgtat 120
cactaaggtc ttcaaagaca tcagctttca aggaggtgcc taagtactgt ttccacacan 180
c 181

<210> 89
<211> 280
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700607713H1

<400> 89
aattgagggg taaacatgtc tttgtgaaat atatgttctt ttacaatact ttgtactaat 60
ttacgtggaa ttattatttg tttctcattg gagatattta ttcagcctca atggcctttc 120
aggaactcct gaatcaggta ggaggcctag ggagattcca gatccttcag atggtttttg 180
ttgtcttcac cagtgttatt gtggtacctc atattataat agagaactta ctgcagccat 240
tcccagtcat cgctgctggg ttcctatcct cgacaatgtc 280

<210> 90
<211> 267
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 6, 77, 128, 143, 146
<223> a or g or c or t, unknown, or other

<220>

<221> misc_feature
<223> Incyte template ID No: 700607873H1

<400> 90
gaaganacac caaagctcac tcactcttca ggttctcact gaacctactg tatcggaagg 60
gacttcacct cccagangtc catttttatg aagactgttg agacagcttt ccagaaacta 120
gaaccatntg gaagatagac ctnggngtat tcctgtgcgg attatcttga ttacgttaat 180
taattctgga tgggactagg ctaaagtgtc atcatgattt tccattaaca aggtgcacag 240
atgctacaaa tggctgggag aaatcct 267

<210> 91
<211> 258
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 11, 109
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700607972H1

<400> 91
gaggctaagc ntgtgcctcc tggtctctct ggggcagtcg cccgcgcgcc aagacctttc 60
gctgacctca gcgtcccgct gctgcgcaag gaagggcggg gccactgcng tctggacagc 120
gtccgaaggc agcgagtcct ctggaggccg ccgtagtgca gaggagtcgg ttgtcacgtg 180
acccaaggtt agaccatggc ttccaccaag ccgctgtctc gcttctggga gtggggcaag 240
aatatcgttt gcgtgggg 258

<210> 92
<211> 276
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700608519H1

<400> 92
caccacactg catctgccct acatgccacc ttacaccatc atctatttcc cctcccgagg 60
tgtggcctca gcgttttcaa gatgaagggc acaagccata gtcactcctg gggaacccct 120
caacactggt gctggaggga tgtggccaag accacgttgg gggcaagacc gagacttggg 180
gcgggactac aattgtggtt ggtggggcca ggactgacct cttagcctcc ataggcagct 240
acactgtctg tcacttttcc tccttccctg tgccgg 276

<210> 93
<211> 295
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure

<222> 35
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700608661H1

<400> 93
cttagtccct gtactctgag ggtaagcctc atcgntcagg atctattgct gctgcttctc 60
cagctggctt ctgtagagac tgacagagac aactttgcag gataaggtag ctatcaagat 120
gctccatttg cgaagttcac agatgctgca gatgttggag agctccttaa ggaaatacct 180
tcctgagtcc ttaaaggttt atgggactgt cttccacatg aaccagggag ccccattcaa 240
gctcaaggct ctggtggaca agtggcctga tttaatacag tggttgtccg tcctc 295

<210> 94
<211> 293
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 16, 179
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700609074H1

<400> 94
ggcgtggagt tggagnagag cgtcaggcgc ctccgggaga agtttcatgg aaaagtgtcc 60
cccaagaagg caggggctct tatgaggaag tttggcagcg accacactgg agttgggcgc 120
tctatcgtgt acgggctcaa gcagaaagat ggacaggagc tgagcaacga tttggacgnc 180
caggacccac cagaggacat gaagcaggac caagatatcc aggcagtagc cacctctctg 240
ttgcccctga cgcaagccaa tcttcgaatg ttccaaagag cccaagatga cct 293

<210> 95
<211> 288
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 66, 138, 152, 236, 242, 252-253, 279
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700609967H1

<400> 95
ctgcctcgca gccccgagcg cgcgcctctc cagctcccgc tccggcttcc ccaaccaggc 60
ttattntggc tcccgacccg gtgcagaccc ctgacccggc ctccgcccag ctccgccaaa 120
tgcgctactt tacttggnag gaggtggcgc angctccggg agggagaagg agcgatggct 180
cgtaatcgac cggaaggtgt acaacatcag cgacttcagt tcgccgccac ccgggnggct 240
cncgggtcat cnnccactag ctggtcagga tgccacggna tcctttgt 288

<210> 96
<211> 164
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700625315H1

<400> 96
gttagagcag ttacactgaa ccaaagtgac tgagtttgta cagacggtaa tccgtaccaa 60
gcacactcac tgtcctgatc tgaacaccca gcaaggttca tgtccgtgct aagtttgcag 120
cattgtgttc ttttgcattc tttttttact tttattaaag gttc 164

<210> 97
<211> 225
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 28, 59, 65, 73, 79, 84, 88
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700627089H1

<400> 97
cgcggcggct gcagcaggcc accatggnag agcttcagga ggtgcagatc actagaggng 60
aagcnattgt tgncaggtnt ggcnttantg gttgagtcta tcactccagt gggttccctg 120
tttgctctgg catcatactc catcatcttc ctcaagcttt tctcctaccg ggatgtcaat 180
ctgtggtgcc gccagcgaag ggtcaaggcc aaagctgtgt ctgca 225

<210> 98
<211> 265
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 62, 264
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No: 700629293H1

<400> 98
atgaccttta acttttctaa aaatgtgaag ttttgtactt atatatatca gctaaagtat 60
tntagcattc tttagtgtac ttagtttgat gccactttta gtgtttttgt tgcttttgtc 120
tgatttttat gaatgttcat tttaagactc cttgttgaaa tgggacagtt tcgttctttg 180
ataagcccga gaagaggatt cccttgggtg ttgacctcct ctgcatgatg tgcccaagca 240
tctgaactgc aaccaaggcc tttnc 265

<210> 99
<211> 95
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No: 700643961H1

<400> 99
agtataacca ggggccatct gaaagttgtt ctctagccag ataagccact atgagcggta 60
agcatcagtg ctaacgcaga aacttcctga gcagc 95

<210> 100
<211> 307
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 8, 20, 21, 23, 47, 302
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700061625H1

<400> 100
gggaaganac actggcttan nanttggttc tgagggaagc tctctgngtg gatacatatt 60
tataggattt gggtgcaacg aactgtgtga ctagttcagt tcaattcagg gagcaggcag 120
aaaccagcag gtattagaag agatgttcta tatacagagt tctgaggcac tgcagattct 180
gaagaattcc ctaaggaagc acctccctga gtccttaaag gttatgggac tgtctccaca 240
tgaaccaggg aaacccattc aagctcaagg ctgtggtgga caagtggctg atttaatact 300
gntgtta 307

<210> 101
<211> 97
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 48, 90
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700062809H1

<400> 101
ccctgagagt ctccaaaatg tttcagtctg ttataagaac cccattanct cactaaagct 60
tgcactggct gactcctgga cggggttccn gaatgtg 97

<210> 102
<211> 214
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700139104H1

<400> 102
gagatgttcc ctgtcatcga acagtatgga gacattttgg taaaatactt gaggcaagag 60
aaaggcaaac ctgtccctgt gaaagaagtg tttggtgcct acagcatgga tgtgatcacc 120
agcacatcat ttggagtgaa tgttgattcc ctcaacaacc cgaaggatcc ttttgtggag 180
aaagccaaga agctcttaag aattgatttt tttg 214

<210> 103
<211> 265
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700139953H1

<400> 103
aggaagccct gcagagcatc agaggcccag ctagagggac aacacagagg agtaatttgc 60
tgacagacct gcagggatgg acctgctttc agctctcaca ctggaaacct gggtcctcct 120
ggcagtcgtc ctggtgctcc tctacggatt tgggacccgc acacatggac ttttcaagaa 180
acaggggatt cctgggccca aacctctgcc tttttttggc actgtgctga attactatat 240
gggtttatgg aaattcgatg tggag 265

<210> 104
<211> 306
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 2, 61, 68, 111, 139, 263, (296)...(298), 305
<221> unsure
<222>
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700289281H1

<400> 104
ancagatccg ggcactggag agggagctgc aagcaggagc agtcaagagt gtggtcagaa 60
ncaccgtnag tggaccagca gggccggctt tttctaccat ggcggcccaa ngctatggct 120
attaccgcac tgtcatatnc acagccatgt tcggaggcta cagcctttac tacttcaacc 180
gcaaaacctt ctcttttgtc atgccctcct tggtggatga gatcgctctg gacaaggacg 240
atttgggggc tcatcacgag canccagtcg gcagcctacg catcagcaag tttgtnnncg 300
gggtnt 306

<210> 105
<211> 183
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700303922H1

<400> 105
ggcagcattg atccttatgt atatctgccc tttggaaatg gacccaggaa ctgcattggc 60
atgaggtttg ctctcatgaa tatgaaactc gctctcacta aagttctgca aaacttctcc 120
ttccagcctt gtaaggaaac acagatacct ctgaaattaa gcagacaagg acttcttcaa 180
cca 183

<210> 106
<211> 290
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 56
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700305783H1

<400> 106
ctaacagtga atacatagct gcgatcctgg aactcagctc cctcatagtg aaacgncaac 60
gccagccctt cctgtacctg gacttcctgt attgcctcac tgctgatggg cggcgcttcc 120
gcaaggcctg cgacgtggtg cacaacttca cagatgctgt catcagggag agacgcagca 180
ccctcaatac ccagggcgtt gatgaattcc taaaggccag ggctaagact aaaactttag 240
actttattga tgttctcttg ctggccaagg atgagcatgg gaaggggctg 290

<210> 107
<211> 177
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700329969H1

<400> 107
gctatatcag gaggggaccc atgctgtgtc cttctgagat ctaacaggat taaccaatat 60
gtaaactaga ggaagtggtt ggcctgcact gggcaagccc tctaggactc catccaagaa 120
agaccagttg gtgttgctct agaggcaaag aaacccataa ggagctggca gtaaaac 177

<210> 108
<211> 188
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 114, 116
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700368493H1

<400> 108
aaagcttccc aatctgtgtg ctcacttggg aggatggcat taggagccag ggctggccat 60
gggtaccttа ctttcctccc tggggtatgc ccaggagaat ggagaaaaaa aaangnttta 120
aagaaaaaat attttaaatt tgatgctggc ctttttcaat tgtattgagt aaaagtgttc 180
aagttgtc 188

<210> 109
<211> 255
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 156, 157
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700376694H1

<400> 109
ttatgattat caattttaca taacattaat attatatcaa acctccttaa gaaaatgagt 60
atggatgttc acagtatgtt tgatttttat ctacaagaat gaatctgatt cagaatgctt 120
ttcagctgac atacagagca ctaaatactt taaggnnaac cataggtctg aatctcttaa 180
gaattctcag tctctatggg atgtagggac gcattataaa tgcattaatc cttatagtca 240
atcctgtgcc tagga 255

<210> 110
<211> 284
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 16, 25, 54, 63, 68, 70, 80, 141, 154, 210, 275
<221> unsure
<222>
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700483803H1

<400> 110
gaggaacgcc gccgcntcgc tcggnatcct acaccaatca ggaagctgct gtcnagccat 60
ggnggganan gagaagccan ctcaagaggc tgacgtggaa cccatggtaa catcaggggc 120

ctcagaagca gtgccaaggg ngctttctgg agancctcag aacatctctg atgtagatgc 180
cttcaacttg ctcctggaga tgaaactgan acgacggcgt gaggtcccaa ccttccatgt 240
actgtgaccc agctagtggc cgaggatggc agcanggtgt atgt 284

<210> 111
<211> 258
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700502844H1

<400> 111
ggcagacctt agcagattct ggatgccatc agtgactgaa gccattaagg agactcggct 60
ggagtagcag ctaagaggac agagagacaa gggctacgag gcagcaatat aaacagatct 120
ggtgttgctg agatttgaga cgaaggtttc ccatggcttc ttttcacatc cgccagttcc 180
aggagaggga ctatgaacag gtcgtggata tgttctccag gggaatgaag gaacacatcc 240
ccactgcctt ccgccact 258

<210> 112
<211> 250
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700503415H1

<400> 112
gtagctttcc ccttttgctg gcacagaagt ctgtccatct gcaagcgctt tggaacacag 60
actgcctgga gccaccttcc tttgggagac cttcctgcct cagctgtcgt cctgtgtcgt 120
cattcactaa agctcctgac gtcagattaa gcaagcagtg atgggttaca ttagagacaa 180
gccgcagaga taaggcctgt tgctgtttcg cagataatga tgagttttaa ttacccactg 240
gtttgtatgg 250

<210> 113
<211> 278
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 40, 44, (53)...(55), 60, 69, 72, 74, 159, 234
<221> unsure
<222>
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700514914H1

<400> 113
ccactgcagc gccccccccc aaagatggaa agattctgcn ttancttcat atnnncttan 60

aacatttgna cncnattctt taactttaga atctccccta gagcctgttc ttctttaaac 120
acccttatg ctggagtagg atgatggctg agtttcttna aaagagctta aatatagagt 180
cacaaacatg agatagatgc ctgccgccca ctctttccac aaactcgaga accnctttgt 240
gcacgcagct gccatggaag gaaatcctgg ggcttctt 278

<210> 114
<211> 308
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte template ID No.: 700519169H2

<400> 114
cccaatcttt taagactttt cataatgatg ttaagaccag ggcagactat ctactggcca 60
tggactgcag ccaggagatg ccctgcccct cctgggacct gcacaccacc tctctgggga 120
acttgacaaa ggtccctaag gctaagggag gtctccttcc tactaggtcc ctgactttga 180
ctctgtggtt ctctaggaac cgtgtgcaca cttgtctctg ttgtaaccac aaagggcagt 240
agcacctagc atgtcatgtc ctgccccggc tgcttgcctc ccacccaccc aggattctct 300
gggctggc 308

<210> 115
<211> 124
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 34, 53, 87, 109, 122
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700535905H1

<400> 115
cctacctctg gtggaaaaag tacatcactc aggngcagct ggtccagttt gtngctgaca 60
atcatccaga ccagctgcgg ggtcatntgg ccgtgctcct tccctctcng gtggctgtac 120
tncc 124

<210> 116
<211> 262
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 148
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700593984H1

<400> 116
gggaaagtgt gttctgaggg ccctgtgggg ccaaggggga ccagcctcac attccacacg 60
tgcgccactc tgcttggagc ctatttattt tgtatttatt tgaacagagt tatgtcctaa 120
ctatttttat agatttgttt aattaatncc ctgtcatttt caagttcatt ttttttattc 180
atatttatgt tcatggttga ttgtacctcc tgtcaccagc tggtggggca ggggagacaa 240
ggtagaaagt ggccacagag tg 262

<210> 117
<211> 267
<212> DNA
<213> Rattus norvegicus

<220>
<221> unsure
<222> 36, 37, 223
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte template ID No.: 700607844H1

<400> 117
ctcgatcgtc cagaccccac ggcgtcacca tgctgnncca tgatcaggaa ctcactgttc 60
gggagcgtgg agacgtggcc ttggcaggtt ttaagcaccg ggggcaagga agacgtctcc 120
tatgaggaaa gagcctgcga agggggaaag tttgctactg tggaagtgac agacaaacct 180
gtggatgagg ctctccggga agcaatgccc aagatcatga agncgtgggg ggcaccaatg 240
acaaaggagt cggcatggga atgacag 267